MicroRNA degradation by a conserved target RNA regulates animal behavior

MicroRNA Degradation by a Conserved Target RNA Regulates Animal Behavior

Angelo Bitetti[1,8], Allison C. Mallory[1,8], Elisabetta Golini[2], Claudia Carrieri[3], Héctor Carreño Gutiérrez[4], Emerald Perlas[5], Yuvia A. Pérez-Rico[1,6], Glauco P. Tocchini-Valentini[2], Anton J. Enright[7], William H. J. Norton[4], Silvia Mandillo[2], Dónal O'Carroll[3], Alena Shkumatava[1,*]

[1] Institut Curie, PSL Research University, CNRS UMR3215, INSERM U934, Paris, France.

[2] Consiglio Nazionale delle Ricerche, Istituto di Biologia Cellulare e Neurobiologia, European Mouse Mutant Archive-Infrafrontier-International Mouse Phenotyping Consortium, I-00015 Monterotondo Scalo (Rome), Italy.

[3] MRC Centre for Regenerative Medicine, Institute for Stem Cell Research, School of Biological Sciences, University of Edinburgh, Edinburgh, UK.

[4] Department of Neuroscience, Psychology and Behaviour, University of Leicester, Leicester, UK.

[5] European Molecular Biology Laboratory (EMBL), Mouse Biology Unit, Monterotondo Scalo, Italy

[6] INSERM, U900, Paris Cedex 05, France.

[7] Department of Pathology, University of Cambridge, Cambridge, United Kingdom.

[8] These authors contributed equally to this work.

[*] Correspondence to: alena.shkumatava@curie.fr

ABSTRACT

Animal microRNAs (miRNAs) repress target transcripts through partial complementarity. By contrast, highly complementary miRNA binding sites within viral and artificially engineered transcripts induce miRNA degradation *in vitro* and in cell lines. Here, we show that a genome-encoded transcript harboring a near-perfect and deeply conserved miRNA binding site for miR-29 controls zebrafish and mouse behavior. This transcript originated in basal vertebrates as a long noncoding RNA and evolved to the protein-coding gene *NREP* in mammals where the miR-29 binding site is located within the 3′UTR. We show that the near-perfect miRNA site selectively triggers miR-29b destabilization through 3′ trimming and restricts its spatial expression in the cerebellum. Genetic disruption of the miR-29 site within mouse *Nrep* results in ectopic expression of cerebellar miR-29b and impaired coordination and motor learning. Thus, we demonstrate an endogenous target RNA-directed miRNA degradation event and its requirement for animal behavior.

INTRODUCTION

MicroRNAs (miRNAs) are ~22 nucleotides RNAs that associate with Argonaute (AGO) proteins to post-transcriptionally repress gene expression[1-3]. Hundreds of miRNAs are expressed in both cell- and tissue-specific patterns throughout the metazoan and plant lineages, impinging on numerous cellular networks[4,5]. Individual miRNAs are predicted to regulate hundreds of target transcripts[6], and thus, the majority of mammalian protein-coding genes appear to be under miRNA regulation[7].

Animal miRNAs recognize the bulk of their target transcripts through limited pairing to miRNA positions 2-8, known as seed pairing[7]. In cases of seed pairing, the miRNA:AGO complex guides target destabilization through shortening of the poly(A) tail and/or repression of translation[7,8]. Although rare in animals, perfect miRNA:target complementarity that includes obligatory pairing to miRNA positions 10 and 11 induces AGO-mediated endonucleolytic cleavage of the target transcript[9-15]. A third class of miRNA target sites defined by almost perfect complementarity along the length of the miRNA but with central mismatches that preclude endonucleolytic cleavage directs target RNA-directed miRNA degradation through 3′ nucleotide removal (trimming) or

additions (tailing)[16-18]. Thus far, artificially engineered sequences containing extensively paired, center-bulged miRNA sites have been shown to trigger miRNA decay by trimming or tailing in mammalian and *Drosophila* cell lines[16,17,19]. Furthermore, by modifying the sequences of these engineered targets to test different degrees of miRNA complementarity, pairing requirements for *in vitro* target RNA-directed miRNA degradation have been defined[16,17]. It has been also proposed that miRNA turnover mediated by these artificial targets is particularly efficient in *in vitro* cultured neurons[17]. In addition, viral sequences containing similar center-bulged miRNA sites can mediate miRNA decay[18,20-22]. While viral transcript-mediated degradation of host miRNAs has been proposed to be advantageous for virus production[18,22], so far the existence of endogenous targets that can trigger miRNA turnover has evaded detection. As such, whether these types of targets are produced *in vivo* and their functional relevance in animals remains unknown.

Here, we identify a deeply conserved and near-perfect miR-29 binding site embedded in a cellular transcript and have elucidated the functional importance of this site *in vivo*. We show that the brain-specific target transcript spatially and quantitatively restricts miR-29b expression by specifically directing miR-29b degradation through 3′ trimming. The failure to destabilize miR-29b in the mouse cerebellum results in multiple brain deficits including impaired balance and motor learning in mice. Moreover, we demonstrate the conserved noncoding function of the target transcript throughout evolution by showing its requirement for the regulation of explorative and anxiety-like behavior in zebrafish. Together, our results suggest that genome-encoded transcripts may have broad *in vivo* relevance by controlling miRNA expression post-transcriptionally through target RNA-directed miRNA degradation.

RESULTS

A Conserved Brain-Enriched Transcript Harbors a Highly Complementary miR-29 Binding Site

As the importance of target RNA-directed miRNA turnover *in vivo* remains unknown, we sought to identify animal transcripts with conserved near-perfect miRNA binding sites and explore their function. We identified a long noncoding RNA (lncRNA) in zebrafish that we called *libra* (*lncRNA involved in behavioral alterations*; previously annotated as linc-*epb4.1l4*[23]) that contains a deeply

conserved miRNA binding site for miR-29, characterized by extensive 5′ and 3′ complementarity with a central 3-nucleotide mismatch (Fig. 1a and Supplementary Fig. 1a, b). The noncoding RNA *libra* shows extensive sequence similarity with the 3′UTR of the mammalian protein-coding gene *NREP* (*Neuronal Regeneration Related Protein*)[23] known to regulate mouse behavior[24] (Fig. 1a). The noncoding sequences spanning several hundreds of nucleotides (Fig. 1a) were detected already at the base of the vertebrate clade and are conserved throughout vertebrate evolution (Fig. 1b). By contrast, the open reading frame (ORF) encoding the 68-amino acid mammalian NREP (also known as P311[25,26]) first appeared in the lobe-finned fishes (e.g. *Coelacanth*; Supplementary Fig. 1c) and was not detected in either of the more ancestral ray-finned or cartilaginous fishes (Fig. 1b and Supplementary Fig.1c and 1d), suggesting that the transcript present in the common ancestor of the vertebrate sequences analyzed was a lncRNA. The deep conservation of the *libra* and *Nrep* noncoding sequences suggests that they may exert important regulatory functions.

To characterize the *libra* and *Nrep* transcripts, we first analyzed and compared their expression levels together with that of the mature miR-29 family members across a set of zebrafish and mouse tissues. Zebrafish *libra* and mouse *Nrep* are almost exclusively expressed in the brain (Fig. 1c and 1e and Supplementary Fig. 1f), whereas miR-29 family members (mouse mmu-miR-29a, b and c and zebrafish dre-miR-29a and b; Supplementary Fig. 1e) are expressed across various tissues but are enriched in both the zebrafish and mouse brains (Fig. 1d and 1e and Supplementary Fig. 1f). Individual mouse miR-29 family members have a similar but non-identical expression pattern in the brain (Fig. 1e), with differential expression patterns observed in the cerebellum. miR-29a and miR-29c are co-expressed with *Nrep* in the granular layer while miR-29b is expressed in Purkinje cells and appears mutually exclusive to the *Nrep* expression domain in the granular layer (Fig. 1e). Notably, the individual members of the miR-29 family have different degrees of complementarity to both *libra* and *Nrep*, with the most extensive predicted pairing being with miR-29b (Supplementary Fig. 1b). Taken together, we show that miR-29b, the most complementary miR-29 family member to *Nrep*, is selectively depleted from the cerebellar granular layer expressing high levels of *Nrep*.

libra **Regulates Explorative and Anxiety-like Behavior in Zebrafish**

The extensive conservation of the sequence and expression pattern between zebrafish *libra* and *Nrep* (Fig. 1a and 1c) raises the possibility that the transcripts impart a biological function through the noncoding elements that have been retained throughout vertebrate evolution (Fig. 1b). Given that *Nrep* regulates mouse behavior by controlling learning, memory and emotional responses[24], we sought to understand if *libra* regulates brain function and generated two loss-of-function alleles in zebrafish. The deletion allele (*libra*^{del}; Supplementary Fig. 2a and 2c) consists of the deletion of nearly the entire *libra* locus. To minimize DNA-dependent effects resulting from the removal of potential *cis* regulatory motifs, we also engineered an inversion allele (*libra*^{inv}) where 5.5 kb of the most conserved part of *libra* are inverted (Supplementary Fig. 2b and 2c). While the *libra* transcript was not detected in *libra*^{del} mutants, the partially inverted transcript was stably expressed in *libra*^{inv} zebrafish, although its levels were lower than those of the wild type *libra* transcript (Supplementary Fig. 2c). We subjected both *libra* mutant lines to a set of standard behavioral tests[27-31]. In the novel tank diving test[30] (Fig. 2a), both bottom dwelling and increased latency to swim into the upper regions of the tank are typical responses of adult zebrafish to a new environment. Both *libra*^{del} and *libra*^{inv} fish showed little sign of bottom dwelling compared to wild type (Fig. 2b-f). Indeed, *libra* mutant fish spent significantly less time at the bottom of the tank, entered top regions of the tank more often, and swam more slowly and for shorter distances than wild type fish (Fig. 2b-f). The altered explorative and anxiety-like behavior revealed by the novel tank diving test[32] was specific, as no major differences between the three genotypes were detected in further experiments testing aggression or novel object boldness[31], apart from an increase in aggression in *libra*^{inv} fish (Supplementary Fig. 3). The behavioral alterations of *libra* mutants in the novel tank test do not appear to be caused by general deficits in their swimming ability as they displayed no difference from wild type fish in swimming velocity or total distance swum in the aggression test (Supplementary Fig. 3d and 3e). Together, our results demonstrate that in basal vertebrates *libra* is required to regulate exploration and anxiety-like behavior.

The *Nrep* miR-29 Site Restricts miR-29b Expression Domain in the Mouse Cerebellum

Next, we sought to examine if the regulatory noncoding function to control animal behavior is retained in the conserved 3'UTR of mammalian *Nrep*. As *Nrep*$^{-/-}$ mice lack the entire orthologous *Nrep* transcript[24], it is not possible to discriminate between the coding and noncoding contributions to overall gene function. We therefore specifically uncoupled *Nrep* function from miR-29

regulation by generating the *Nrep* miR-29 scrambled allele (*Nrep*$^{miR\text{-}29scr}$), where nine point substitutions were introduced in the miR-29 binding site to disrupt both the seed pairing and the 3' complementarity (Fig. 3a and Supplementary Fig. 4a). Mice homozygous for the *Nrep*$^{miR\text{-}29scr}$ allele, hereafter referred to as *Nrep*$^{miR\text{-}29scr}$ mice, are viable and show normal brain morphology (Supplementary Fig. 4b). Importantly, scrambling the miR-29 site did not impact the expression of the *Nrep* transcript (Fig. 3b and Supplementary Fig. 4c). Strikingly, we found that miR-29b, which is normally confined to the cerebellar Purkinje cells (Fig. 1e), gained an additional expression domain within the cerebellar granule cell layer of *Nrep*$^{miR\text{-}29scr}$ mice (Fig. 3c, middle panel). The expanded miR-29b expression domain suggests that *Nrep* normally prevents miR-29b from accumulating in the granular layer. No changes in the miR-29a or miR-29c expression pattern were detected in the brains of *Nrep*$^{miR\text{-}29scr}$ mice (Fig. 3c). Our results demonstrate that the miR-29 target site in the *Nrep* 3'UTR uniquely regulates the spatial expression of miR-29b.

Failure to Downregulate miR-29 in *Nrep*$^{miR\text{-}29scr}$ Mice Results in Multiple Behavioral Deficits

While the expanded expression of miR-29b in the granular layer of *Nrep*$^{miR\text{-}29scr}$ mice does not appear to affect overall granule or Purkinje cell morphology (Supplementary Fig. 4b), we tested if uncoupling *Nrep* from miR-29 regulation resulted in behavioral alterations associated with granule cell function. Because cerebellar granule cells contribute to motor learning and motor behaviour[33-36], *Nrep*$^{miR\text{-}29scr}$ mice were examined in the rotarod test commonly used to assess normal cerebellar function[37,38]. On day one, no significant coordination deficits were observed. However, *Nrep*$^{miR\text{-}29scr}$ mice showed a significant impairment in balance and motor coordination from day two, falling more quickly from the rotating rod than their wild type littermates (Fig. 4a). Additionally, *Nrep*$^{miR\text{-}29scr}$ mice displayed pronounced motor learning deficits as shown by the lack of improved rotarod performance over the course of four consecutive training days (RM ANOVA, Genotype X Day: $F_{(3,42)} = 5.97$, $P = 0.0017$; Fig. 4a and Supplementary Table 1). In summary, the rotarod test showed an abnormal cerebellum-associated behavior in *Nrep*$^{miR\text{-}29scr}$ mice.

Because *NREP* controls emotional responses, as well as learning and memory[24], we tested if *Nrep*$^{miR\text{-}29scr}$ mice have additional behavioral deficits and profiled them for spatial learning, fear conditioning and anxiety-related behavior[39,40]. Similar to *Nrep*$^{-/-}$ mice[24], *Nrep*$^{miR\text{-}29scr}$ animals showed reduced contextual learning in the fear conditioning test (RM ANOVA, Genotype x Time

(before and 24hr after shock): $F_{(1,24)} = 6.13$, P=0.0208; Fig. 4b, Supplementary Fig. 5a and 5d, Supplementary Table 1) and reduced swimming velocity in the Morris water maze test (RM ANOVA, Genotype: $F_{(1,15)} = 12.76$, $P = 0.0028$; Fig. 4d, Supplementary Table 1). In addition, no alteration in anxiety-related behavior in the elevated plus maze test was revealed in *Nrep*$^{miR\text{-}29scr}$ animals (Supplementary Fig. 5f and 5g, Supplementary Table 1). In contrast to *Nrep*$^{-/-}$ mice[24], we detected no differences in cue-dependent learning and emotional memory in the fear conditioning test (Fig. 4c, Supplementary Fig. 5e and Supplementary Table 1) or spatial learning in the Morris water maze (Fig. 4e and 4f, Supplementary Table 1) in *Nrep*$^{miR\text{-}29scr}$ mice. Notably, the motor learning deficits of *Nrep*$^{miR\text{-}29scr}$ mice detected in the rotarod test (Fig. 4a) appear to be brain malfunctions rather than movement defects associated with muscle deficits as *Nrep*$^{miR\text{-}29scr}$ mice were not impaired in their general activity before shock in the fear conditioning test (Supplementary Fig. 5b, Supplementary Table 1). Additionally, the frequency of total entries in the elevated plus maze test (Supplementary Fig. 5f), and the latency and velocity to reach the visible platform in the Morris water maze (Fig. 4d and 4e) indicated normal locomotion and swimming abilities. Nevertheless, it is possible that the balance and motor coordination impairments of mutant mice detected in the rotarod test (Fig. 4a) may be accompanied by neuromuscular or other defects (e.g. altered grip strength or tactile sensitivity). Taken together, we show that scrambling the miR-29 site within the *Nrep* transcript leads to an ectopic miR-29b expression domain in the cerebellum and the impairment of several brain functions resulting in behavioral deficits.

The miR-29 Site of *Nrep* Directs miR-29b Degradation through 3′ Trimming

The complementarity of miR-29b to *Nrep* (Supplementary Fig. 1b) would make it a likely substrate for target RNA-directed miRNA degradation[16,17] and the possible basis for the regulated expression in the granular layer of the cerebellum. Because we could not isolate a pure granular layer cell population that specifically expresses *Nrep,* we turned to an *in vitro* cellular system to explore the mechanism by which miR-29b is regulated. The *Nrep*$^{miR\text{-}29scr}$ allele was introduced into mouse embryonic stem cells (ESCs), and these cells were subsequently differentiated into neural progenitor cells (NPCs)[41,42] that, in contrast to undifferentiated ESCs, express abundant levels of *Nrep* (Fig. 5a and Supplementary Fig. 6a). No differences in cell morphology or in the expression of the pluripotency marker *Oct4* or the neuronal marker *Nestin* (Supplementary Fig. 6b) were observed between wild type and *Nrep*$^{miR\text{-}29scr}$ ESCs and NPCs, respectively. As detected in the

cerebellum of *Nrep*$^{miR\text{-}29scr}$ mice, the expression level of the *Nrep* transcript in NPCs was not significantly altered upon scrambling the miR-29 site (Fig. 5a and Supplementary Fig. 6c); however, mature miR-29b accumulated to significantly elevated levels in *Nrep*$^{miR\text{-}29scr}$ NPCs (Fig. 5a and 5b). The levels of mature miR-29a also were increased in *Nrep*$^{miR\text{-}29scr}$ NPCs (Fig. 5a and 5b), although to a far lesser extent than for miR-29b. To exclude the contribution of transcriptional regulation or miRNA processing, we measured the levels of both miR-29 primary transcripts (pri-miR-29s) as well as miR-29s* (passenger strands), and found that neither showed significant changes upon disruption of the *Nrep* miR-29 site (Supplementary Fig. 6d and 6e). These results indicate that post-transcriptional and post-miRNA biogenesis mechanisms regulate mature miR-29 levels in *Nrep*$^{miR\text{-}29scr}$ NPCs. Indeed, small RNA sequencing to identify miR-29 isoforms produced through 3′ trimming or tailing revealed that scrambling the *Nrep* miR-29 site specifically prevented mature miR-29b from being 3′ trimmed (Fig. 5c). Trimming was not detected for miR-29a or miR-29c (Fig. 5c), nor was it detected for the five most abundant miRNAs expressed in either wild type or *Nrep*$^{miR\text{-}29scr}$ NPCs (Supplementary Fig. 6f). Moreover, non-genome-templated 3′ nucleotide additions (tailing) to miR-29a, miR-29b or miR-29c were not observed in wild type or *Nrep*$^{miR\text{-}29scr}$ NPCs (Fig. 5c). These results directly indicate that target RNA-directed miRNA degradation through trimming is the mechanism by which *Nrep* suppresses miR-29b, demonstrating that the noncoding sequence of *Nrep* exerts a key regulatory function.

DISCUSSION

Here, we report the first example of a target RNA-directed miRNA degradation event in animals and demonstrate its physiological relevance *in vivo*. Intriguingly, in zebrafish, the near-perfect miR-29 binding site is located in a long noncoding RNA, whereas in mice this miRNA target site is located within the 3'UTR of an ORF-encoding transcript. Our findings indicate that the mammalian *Nrep* protein-coding gene has in fact evolved from an ancestral noncoding transcript and has retained the regulatory noncoding function in the 3′UTR. Disruption of the *Nrep* miR-29 site revealed that *Nrep*-mediated suppression of miR-29b in the cerebellar granular layer is required for normal motor learning and balance in mice, although it is possible that additional neuromuscular features could be implicated. Furthermore, we found a partial overlap in the behavioral deficits observed in *Nrep*$^{-/-}$ mice that lack the entire *Nrep* transcript[24] and *Nrep*$^{miR\text{-}29scr}$

mice, confirming the importance of *Nrep*'s noncoding sequences. Moreover, all 20 bases paired between *libra* (*Nrep*) and miR-29b are conserved in all of the examined vertebrate genomes (Fig. 1a and Supplementary Fig. 1a). Although the zebrafish *libra* alleles presented in this study do not allow the specific uncoupling of *libra* from miR-29 regulation, the deep sequence conservation of the miR-29 site together with the functional role of *libra* in zebrafish behavior indicate that the mechanism of target RNA-dependent miRNA decay likely also exists in more distal vertebrates.

The sequence conservation between the ancestral noncoding RNA *libra* and the mammalian *Nrep* 3′UTR extends several hundreds of nucleotides beyond the miR-29 site (Fig. 1a). Although the precise function of these flanking sequences is unknown, it is reasonable to hypothesize that they could have important roles in supporting miR-29 turnover or could contribute to other independent functions of the transcript. The extended sequence conservation also raises the question of whether the miR-29 complementary sequence within the *libra* and *Nrep* transcripts might have occurred independently of the miRNA as has been proposed for some miRNA sites[43]. This situation does not appear to be the case for the *libra* or *Nrep* transcript for two reasons. Firstly, in all vertebrate genomes in which *libra* conserved sequences could be detected, miR-29 genes were also found. Secondly, miR-29 likely has a more ancestral origin than the *libra* and *Nrep* loci because miR-29 genes but not *libra* or *Nrep* loci were detectable in non-vertebrate animals such as *Ciona* and sea urchin (data not shown). In summary, although extended conservation is observed in the *libra* and *Nrep* transcript, our results clearly pinpoint the miR-29 site located within this conserved sequence block as a functionally important sequence element.

Precise miR-29 dosage is critical in the brain as deletion of only one of the two miR-29 gene clusters (the miR-29a/b-1 cluster) leads to ataxia and cerebellar alterations in mice[44]. In this study, we show that upon the disruption of the *Nrep* miR-29 site, the expanded expression of miR-29b into the cerebellar granular layer results in multiple brain malfunctions. Interestingly, while miR-29 is broadly expressed in the brain, the target transcript *Nrep* is almost exclusively confined to the cerebellar granular layer and specifically delineates the spatial expression of mature miR-29b in the cerebellum. Thus, the post-transcriptional degradation of a broadly expressed miRNA mediated by a cell-specific target transcript appears to be an effective mechanism to achieve precise spatial miRNA regulation. miR-29b is expressed from both bicistronic miR-29 clusters, making it difficult

to instruct transcriptional regulation of miR-29b without affecting miR-29a or miR-29c dosage. This restriction could necessitate the selective *Nrep*-mediated post-transcriptional regulation of miR-29b.

As miRNAs are often transcribed in clusters from multiple genomic loci and are generally highly stable[45-48], miRNA destabilization through endogenous targets emerges as an effective post-transcriptional mechanism for the selective regulation of specific miRNAs. Originally proposed through *in vitro* studies[16,17], it has remained unclear until now if miRNA turnover by endogenous targets occurs in animals. *In vivo* evidence for endogenous target RNA-directed miRNA degradation has likely been hampered by the fact that highly complementary, center-bulged miRNA binding sites within cellular transcripts are challenging to predict computationally. Our study suggests that genome-encoded miRNA target sites that direct miRNA turnover represent an exacting mode of miRNA regulation that may be commonly found throughout the noncoding transcriptome of vertebrates.

ACKNOWLEDGMENTS

We thank R. Pillai and E. Heard for comments on this manuscript. We also thank all members of the Shkumatava lab for useful discussions and B. Foret and A-V. Gendrel of the E. Heard lab for advice on ESC differentiation. This work was supported by grants from ERC (FLAME-337440), ATIP-Avenir, La Fondation Bettencourt Schueller and La Fondation pour la Recherche Médicale (FRM DBI201312285578) to A.S. The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (*FP7/2007-2013)* / ERC *grant agreement* n° GA 310206 to D.O'C; and the European Research Council under the European Union's Seventh Framework Programme (*FP7/2007-2013)* grant agreement no. 602805 to W.H.J.N. High-throughput sequencing has been performed by the ICGex NGS platform of the Institut Curie supported by the grants ANR-10-EQPX-03 (Equipex) and ANR-10-INBS-09-08 (France Génomique Consortium) from the Agence Nationale de la Recherche ("Investissements d'Avenir" program), by the Canceropole Ile-de-France and by the SiRIC-Curie program - SiRIC Grant « INCa-DGOS- 4654.

AUTHOR CONTRIBUTIONS

A.B. and A.C.M. contributed to the design, execution and analysis of most of the experiments. C.C. contributed to the design, generation and execution of the *Nrep*$^{miR\text{-}29scr}$ allele. E.G. and S.M. performed and analyzed mouse behavioral experiments. H.C.G., W.H.J.N. and A.B. performed and analyzed zebrafish behavioral experiments. E.P. performed the *in situ* hybridization stainings. Y.A.P-R. performed the phylogenetic and sequence conservation analyses. G.P.T-V. and D.O'C. thought of the *Nrep*$^{miR\text{-}29scr}$ allele. A.J.E. performed the bioinformatic analyses of small RNA sequencing. A.S. conceived the study. D.O'C. and A.S. supervised the study. A.S. and A.C.M. wrote the final version of the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

REFERENCES

1. Bartel, D.P. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116**, 281-97 (2004).

2. Kim, V.N., Han, J. & Siomi, M.C. Biogenesis of small RNAs in animals. *Nat Rev Mol Cell Biol* **10**, 126-39 (2009).
3. Meister, G. Argonaute proteins: functional insights and emerging roles. *Nat Rev Genet* **14**, 447-59 (2013).
4. Carrington, J.C. & Ambros, V. Role of microRNAs in plant and animal development. *Science* **301**, 336-8 (2003).
5. Gurtan, A.M. & Sharp, P.A. The role of miRNAs in regulating gene expression networks. *J Mol Biol* **425**, 3582-600 (2013).
6. Friedman, R.C., Farh, K.K., Burge, C.B. & Bartel, D.P. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res* **19**, 92-105 (2009).
7. Bartel, D.P. MicroRNAs: target recognition and regulatory functions. *Cell* **136**, 215-33 (2009).
8. Filipowicz, W., Bhattacharyya, S.N. & Sonenberg, N. Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nat Rev Genet* **9**, 102-14 (2008).
9. Davis, E. et al. RNAi-mediated allelic trans-interaction at the imprinted Rtl1/Peg11 locus. *Curr Biol* **15**, 743-9 (2005).
10. Yekta, S., Shih, I.H. & Bartel, D.P. MicroRNA-directed cleavage of HOXB8 mRNA. *Science* **304**, 594-6 (2004).
11. Moran, Y. et al. Cnidarian microRNAs frequently regulate targets by cleavage. *Genome Res* **24**, 651-63 (2014).
12. Shin, C. et al. Expanding the microRNA targeting code: functional sites with centered pairing. *Mol Cell* **38**, 789-802 (2010).
13. Karginov, F.V. et al. Diverse endonucleolytic cleavage sites in the mammalian transcriptome depend upon microRNAs, Drosha, and additional nucleases. *Mol Cell* **38**, 781-8 (2010).
14. Park, J.H. et al. Degradome sequencing reveals an endogenous microRNA target in C. elegans. *FEBS Lett* **587**, 964-9 (2013).
15. Hutvagner, G. & Zamore, P.D. A microRNA in a multiple-turnover RNAi enzyme complex. *Science* **297**, 2056-60 (2002).
16. Ameres, S.L. et al. Target RNA-directed trimming and tailing of small silencing RNAs. *Science* **328**, 1534-9 (2010).
17. de la Mata, M. et al. Potent degradation of neuronal miRNAs induced by highly complementary targets. *EMBO Rep* **16**, 500-11 (2015).
18. Marcinowski, L. et al. Degradation of cellular mir-27 by a novel, highly abundant viral transcript is important for efficient virus replication in vivo. *PLoS Pathog* **8**, e1002510 (2012).
19. Baccarini, A. et al. Kinetic analysis reveals the fate of a microRNA following target regulation in mammalian cells. *Curr Biol* **21**, 369-76 (2011).
20. Cazalla, D., Yario, T. & Steitz, J.A. Down-regulation of a host microRNA by a Herpesvirus saimiri noncoding RNA. *Science* **328**, 1563-6 (2010).
21. Libri, V. et al. Murine cytomegalovirus encodes a miR-27 inhibitor disguised as a target. *Proc Natl Acad Sci U S A* **109**, 279-84 (2012).
22. Lee, S. et al. Selective degradation of host MicroRNAs by an intergenic HCMV noncoding RNA accelerates virus production. *Cell Host Microbe* **13**, 678-90 (2013).
23. Ulitsky, I., Shkumatava, A., Jan, C.H., Sive, H. & Bartel, D.P. Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution. *Cell* **147**, 1537-50 (2011).

24. Taylor, G.A. et al. Behavioral characterization of P311 knockout mice. *Genes Brain Behav* **7**, 786-95 (2008).
25. Taylor, G.A., Hudson, E., Resau, J.H. & Vande Woude, G.F. Regulation of P311 expression by Met-hepatocyte growth factor/scatter factor and the ubiquitin/proteasome system. *J Biol Chem* **275**, 4215-9 (2000).
26. Studler, J.M., Glowinski, J. & Levi-Strauss, M. An abundant mRNA of the embryonic brain persists at a high level in cerebellum, hippocampus and olfactory bulb during adulthood. *Eur J Neurosci* **5**, 614-23 (1993).
27. Norton, W. & Bally-Cuif, L. Adult zebrafish as a model organism for behavioural genetics. *BMC Neurosci* **11**, 90 (2010).
28. Stewart, A.M. et al. Molecular psychiatry of zebrafish. *Mol Psychiatry* **20**, 2-17 (2015).
29. Chodroff, R.A. et al. Long noncoding RNA genes: conservation of sequence and brain expression among diverse amniotes. *Genome Biol* **11**, R72 (2010).
30. Bencan, Z., Sledge, D. & Levin, E.D. Buspirone, chlordiazepoxide and diazepam effects in a zebrafish model of anxiety. *Pharmacol Biochem Behav* **94**, 75-80 (2009).
31. Norton, W.H. et al. Modulation of Fgfr1a signaling in zebrafish reveals a genetic basis for the aggression-boldness syndrome. *J Neurosci* **31**, 13796-807 (2011).
32. Egan, R.J. et al. Understanding behavioral and physiological phenotypes of stress and anxiety in zebrafish. *Behav Brain Res* **205**, 38-44 (2009).
33. Kadotani, H. et al. Motor discoordination results from combined gene disruption of the NMDA receptor NR2A and NR2C subunits, but not from single disruption of the NR2A or NR2C subunit. *J Neurosci* **16**, 7859-67 (1996).
34. Galliano, E. et al. Silencing the majority of cerebellar granule cells uncovers their essential role in motor learning and consolidation. *Cell Rep* **3**, 1239-51 (2013).
35. Watanabe, D. et al. Ablation of cerebellar Golgi cells disrupts synaptic integration involving GABA inhibition and NMDA receptor activation in motor coordination. *Cell* **95**, 17-27 (1998).
36. Ceccarelli, M. et al. Altered cerebellum development and impaired motor coordination in mice lacking the Btg1 gene: Involvement of cyclin D1. *Dev Biol* **408**, 109-25 (2015).
37. Crawley, J.N. Behavioral phenotyping of transgenic and knockout mice: experimental design and evaluation of general health, sensory functions, motor abilities, and specific behavioral tests. *Brain Res* **835**, 18-26 (1999).
38. Marazziti, D. et al. Precocious cerebellum development and improved motor functions in mice lacking the astrocyte cilium-, patched 1-associated Gpr37l1 receptor. *Proc Natl Acad Sci U S A* **110**, 16486-91 (2013).
39. Crawley, J.N. & Paylor, R. A proposed test battery and constellations of specific behavioral paradigms to investigate the behavioral phenotypes of transgenic and knockout mice. *Horm Behav* **31**, 197-211 (1997).
40. Mandillo, S. et al. Mice lacking the Parkinson's related GPR37/PAEL receptor show non-motor behavioral phenotypes: age and gender effect. *Genes Brain Behav* **12**, 465-77 (2013).
41. Conti, L. et al. Niche-independent symmetrical self-renewal of a mammalian tissue stem cell. *PLoS Biol* **3**, e283 (2005).
42. Splinter, E. et al. The inactive X chromosome adopts a unique three-dimensional conformation that is dependent on Xist RNA. *Genes Dev* **25**, 1371-83 (2011).
43. Pinzon, N. et al. microRNA target prediction programs predict many false positives. *Genome Res* **27**, 234-245 (2017).

44. Papadopoulou, A.S. et al. Deficiency of the miR-29a/b-1 cluster leads to ataxic features and cerebellar alterations in mice. *Neurobiol Dis* **73**, 275-88 (2015).
45. Gatfield, D. et al. Integration of microRNA miR-122 in hepatic circadian gene expression. *Genes Dev* **23**, 1313-26 (2009).
46. Marzi, M.J. et al. Degradation dynamics of microRNAs revealed by a novel pulse-chase approach. *Genome Res* **26**, 554-65 (2016).
47. Rissland, O.S., Hong, S.J. & Bartel, D.P. MicroRNA destabilization enables dynamic regulation of the miR-16 family in response to cell-cycle changes. *Mol Cell* **43**, 993-1004 (2011).

FIGURE LEGENDS

Figure 1 Evolution and conserved expression of the *libra* and *Nrep* transcripts.
(a) The *libra* locus in zebrafish, and the *NREP* loci in mouse and human. Boxed gray areas and PhastCons plots based on the 8-genome alignment indicate the location of deeply conserved sequences. The PhastCons plot is relative to the zebrafish locus. Dark blue boxes in the mammalian loci represent the NREP-coding ORF, light blue boxes represent 5′ and 3′ UTRs. The consensus sequence logo shows conservation of the miR-29 site, with vertical lines indicating Watson–Crick and wobble pairing. (b) Cladogram representing the Bayesian phylogeny inferred from the most 3′ region of sequence conservation of the *libra* and *NREP* genes in 25 vertebrates. The presence "+" or absence "-" of the NREP ORF and the three conserved noncoding sequence blocks (gray boxes in panel a) are indicated for each species. Partial ORF sequences are denoted with "+/-". (c) Zebrafish *libra* (left) and mouse *Nrep* (right) expression across adult tissues and in mouse embryonic fibroblasts (MEFs) detected by RNA blots. *18s* rRNA was used as a loading reference. (d) Expression of the miR-29 family members in adult zebrafish (left) and mouse (right) tissues and cells. *U6* RNA was used as a loading reference. Uncropped blot images are shown in Supplementary Data Set 1. (e) *Nrep* and miR-29a, miR-29b and miR-29c expression by *in situ* hybridization on adult mouse brain sections. Representative sections for each probe are shown (wild type animals n=4). Left panel, whole brain sections, the cerebellum is outlined by a gray box; right panel, a zoom-in of the cerebellum. Scale bars represent 100 μm. Arrowheads point to Purkinje cells; granular layer (GL). Control experiments using scrambled miRNA probes were also performed (Supplementary Fig. 1g and 1h).

Figure 2 *libra* regulates explorative behavior in zebrafish.
(a) In the novel tank diving test, the tank was divided virtually into three sections to demark the exploratory activity of zebrafish. The movement of individual fish was recorded for five minutes. (b) Representative swimming traces (red) in the novel tank diving test of wild type, $libra^{del}$ and $libra^{inv}$ zebrafish analyzed with video-tracking software. Each trace represents an individual adult fish from one representative experiment; wild type n=14, $libra^{del}$ n=15, $libra^{inv}$=14. (c) Time in seconds (s) spent at the bottom of the tank (d) number (n) of top entries (e) total distance in centimeters (cm) swum and (f) velocity (cm/s) of wild type and *libra* mutant zebrafish in the novel tank diving test. Data are presented as mean ± SEM, One-way ANOVA test followed by Dunnett's multiple comparisons test: $^{**}P < 0.01$, $^{***}P < 0.001$. Each dot represents an individual adult fish from one representative experiment; wild type animals n=14, $libra^{del}$ animals n=15, $libra^{inv}$ animals n=14. The novel tank test was performed in three independent experiments. Detailed statistical analyses and source data for panels c-f are provided in Supplementary Table 1 and Supplementary Data Set 2.

Figure 3 Disruption of the *Nrep* miR-29 site leads to expanded miR-29b expression in cerebellum.
(a) Schematic of DNA point substitutions at the miR-29 site of mouse *Nrep* (mutated nucleotides shown in red) disrupting both the miR-29 seed and 3′ complementarity pairing. Introducing

substitutions at the miR-29 site did not generate a new miRNA binding site according to miRanda v3.0[49] against the entire set of miRNAs available for *Mus musculus* (miRBase 21). (b) *Nrep* expression by *in situ* hybridization on brain sections in wild type (left) and miR-29 scrambled site mutant (right) mice. Upper panel, whole brain; bottom panel, zoom-in of the cerebellum. Scale bars represent 100 μm. One representative section for each probe is shown (wild type animals n=4, $Nrep^{miR\text{-}29scr}$ animals n=4). (c) Expression of individual miR-29 family members in wild type (left) and miR-29 scrambled mutant mouse cerebellum (right) by *in situ* hybridization on brain sections. One representative section for each probe is shown (wild type animals n=4, $Nrep^{miR\text{-}29scr}$ animals n=4). Arrowheads point to Purkinje cells; granular layer (GL). Scale bars represent 100 μm.

Figure 4 $Nrep^{miR\text{-}29scr}$ mice display multiple behavioral deficits.
(a) Motor coordination and motor learning of $Nrep^{miR\text{-}29scr}$ (purple) and wild type (blue) mice measured as the latency to fall (seconds, s) in the rotarod test over the course of four days; wild type (WT) animals n=9, $Nrep^{miR\text{-}29scr}$ animals n=7. For all panels, data are presented as individual data points and mean ± SEM; Bonferroni *post hoc* tests: WT *vs* $Nrep^{miR\text{-}29scr}$ $*P < 0.05$, $**P < 0.01$, $***P < 0.001$, (ns) not significant. Day split by Genotype: for WT, day 1 *vs* day 2 $P < 0.01$, day 1 *vs* day 3 $P < 0.0001$, day 1 *vs* day 4 $P < 0.0001$; for $Nrep^{miR\text{-}29scr}$ day 1 *vs* day 2 ns, day 1 *vs* day 3 ns, day 1 *vs* day 4 ns. (b) Percentage of freezing of $Nrep^{miR\text{-}29scr}$ (purple) and wild type (blue) mice in the fear conditioning test before shock (min 3-4, conditioning session) and after shock (6 min total, context session). Bonferroni *post hoc* tests: WT *vs* $Nrep^{miR\text{-}29scr}$ $*P < 0.05$; Wild type: baseline *vs* context $P < 0.0001$; $Nrep^{miR\text{-}29scr}$ baseline *vs* context $P < 0.001$. (c) Percentage of freezing of $Nrep^{miR\text{-}29scr}$ (purple) and wild type (blue) mice in the fear conditioning (cue session). For panels b and c, wild type animals n=14, $Nrep^{miR\text{-}29scr}$ animals n=12. (d) Swim velocity (cm/s), and (e) latency to reach platform (s) in the Morris water maze test. Bonferroni *post hoc* tests: For wild type, day 2 *vs* day 3 not significant (ns), day 2 *vs* day 4 $P < 0.001$, day 2 *vs* day 5 $P < 0.0001$; for $Nrep^{miR\text{-}29scr}$, day 2 *vs* day 3 ns, day 2 *vs* day 4 $P < 0.01$, and day 2 *vs* day 5 $P < 0.05$. (f) Percentage of time spent in each of the four quadrants in the probe trial of the Morris water maze test. Bonferroni *post hoc* tests: For wild type, $P < 0.0001$ correct *vs* right and opposite, $P < 0.01$ correct *vs* left; for $Nrep^{miR\text{-}29scr}$, $P < 0.0001$ correct *vs* right, opposite and left. Wild type animals n=8, $Nrep^{miR\text{-}29scr}$ animals n=9. Each dot represents an individual mouse. Detailed statistical analyses and source data are provided in Supplementary Table 1 and Supplementary Data Set 2.

Figure 5 The near-perfect *Nrep* miR-29 site triggers miR-29b degradation by 3′ trimming.
(a) *Nrep* and miR-29a, b and c expression in wild type and $Nrep^{miR\text{-}29scr}$ NPCs detected by RNA blots. Three biological replicates are shown for each genotype. Uncropped blot images are shown in Supplementary Data Set 1. (b) qRT-PCR measuring miR-29a, miR-29b and miR-29c levels in the same wild type (blue) and $Nrep^{miR\text{-}29scr}$ (purple) NPC lines as in panel a. Each dot represents an individual biological replicate NPC population. qRT-PCRs for each biological replicate were performed in technical triplicate. *U6* was used as a reference gene. Data are presented as mean ± SEM; unpaired t-tests: $*P < 0.05$, $**P < 0.01$; (ns) not significant at 95% confidence level. (c) The fraction of small RNA sequencing reads with coverage of at least *n* indicated nucleotides (nt) along

the length of miR-29a, miR-29b and miR-29c. For each miRNA, dashed lines delineate the canonical length and the regions tested for trimming and tailing. Five nucleotides either side of the canonical length were used to test for significant differences in the mean values over these positions using a paired t-test for difference of means. Three biological replicates are shown for each genotype. $^{**}P < 0.01$; (ns) not significant at 95% confidence level. Detailed statistical analyses and source data are provided in Supplementary Table 1 and Supplementary Data Set 2.

ONLINE METHODS

Analyses and source data are provided in Supplementary Table 1 and Supplementary Data Set 2.

Animal models

Generation of CRISPR/Cas9–mediated zebrafish mutants

Two independent zebrafish mutants of *libra* (previously annotated as *linc-epb4.1l4*[23]), *libra*^{del} and *libra*^{inv}, were generated using CRISPR/Cas9-mediated genome editing. Briefly, 9 ng of each sgRNA (Supplementary Fig. 2 and Supplementary Table 2) and 150 ng of *Cas9* mRNA were co-injected into one-cell stage AB zebrafish embryos[50]. sgRNAs and *Cas9* mRNA were generated as described previously[51] using the codon optimized plasmid JDS246 for the *Cas9* mRNA synthesis (Addgene #43861). All of the RNAs were purified with the RNeasy Mini Kit (Qiagen). Genomic DNA was extracted in 1× TE and 2 µg/µl proteinase K (Roche) at 55°C for 3 hours then 94°C for 10 minutes from either embryos or adult tissues and directly used for PCR genotyping, mapping and DNA sequencing using the primers listed in the Supplementary Table 2. All zebrafish were bred and maintained at the Institut Curie, Paris in accordance with the current European Directive 2010/63. Zebrafish were staged using standard procedures[52].

CRISPR/Cas9–mediated homologous recombination in mice

The C57BL/6N mouse strain was used to generate the *Nrep*$^{miR\text{-}29scr}$ allele by CRISPR/Cas9 gene editing[53]. The scrambled miR-29 binding site was introduced into the 3′ UTR of *Nrep* by using two sgRNAs designed in the fifth exon (Supplementary Fig. 4a), Guide 4 (Supplementary Table 2) was designed internal to the miRNA seed site; Guide 5 (Supplementary Table 2) was designed 43 bp further downstream. A 200 nt single stranded DNA oligo (ssDNA *Nrep*$^{miR\text{-}29scr}$; Supplementary Table 2) with 90 bp homology arms flanking both sides of the scrambled miRNA seed site was designed and manufactured by Ultramer (IDT). Two-pronuclei stage mouse embryos were injected with 25 ng Cas9 mRNA, 12.5 ng of sgRNA and 12.5 ng of donor oligo. The presence of the scrambled sequence was scored by PCR using the genotyping primers listed in Supplementary Table 2 followed by *BamH1* digestion of the amplified DNA region. *Nrep*$^{miR\text{-}29scr}$ animals were backcrossed to WT C57BL/6N for 8-12 generations.

All mice were bred and maintained in EMBL Mouse Biology Unit, Monterotondo in accordance with current Italian legislation (D.Lgs 26/2014) under license from the Italian health ministry.

Sequence analyses and phylogenetic tree reconstruction

Predicted sequences with high similarity to the human *NREP* loci from representative vertebrate species of diverse classes were collected from the available genomes in UCSC (https://genome.ucsc.edu/cgi-bin/hgGateway, last accessed July 19, 2016). Sequences were queried using the protein and DNA sequences from human via BLAST to identify NREP protein sequences (Supplementary Note 1) and *libra* and *NREP* DNA sequences. The most 3' block of conserved DNA sequence was chosen for phylogenetic analysis because it was the block with the highest sequence variability among the three conserved sequence blocks (Fig. 1a and Supplementary Note 2). Protein and DNA sequences were aligned using Clustal Omega with default parameters[54-56]. Phylogenetic reconstruction was performed with MrBayes v3.2.6 using the DNA multiple alignment, an evolutionary model of equal rates of substitution (F81) and four simultaneous runs. After five million generations, tree and branch length information was summarized discarding the first 25% of sampled trees. The consensus tree was visualized using FigTree (http//tree.bio.ed.ac.uk/software/figtree).

Sequences corresponding to the miR-29 binding site in the *libra* and *NREP* transcripts from 30 vertebrate species (Supplementary Note 3) were aligned using Clustal Omega with default parameters. This alignment was used to generate a sequence logo with WebLogo 3.5.0[57].

mRNA/lncRNA gel blot analyses

Total RNA was isolated using TRIzol (Invitrogen) according to manufacturer's instructions, separated on 1% agarose gels containing 0.8% formaldehyde and transferred to nylon membrane (Nytran SPC, GE Healthcare) by capillary action. Blots were hybridized with α-UTP ^{32}P-labeled RNA probes at 68°C in ULTRAhyb buffer (Ambion) as recommended by the manufacturer. RNA probe templates were amplified from adult mouse and zebrafish brain cDNA by PCR using the primers listed in Supplementary Table 2 (the sequence of the T7 promoter is underlined) and transcribed *in vitro* (RNA MaxiScript, Ambion) in the presence of α-UTP ^{32}P. Both mouse and zebrafish RNA gel blots and hybridizations in Fig. 1c were performed in technical duplicates. The gel blots and hybridizations in Fig. 5a and Supplementary Fig. 6a were performed in technical duplicate.

Small RNA blot analyses

Total RNA was isolated using TRIzol (Invitrogen), separated on a 15% denaturing polyacrylamide gel, electroblotted to nylon membrane (Hybond-NX, GE Healthcare), chemically cross-linked[58] and hybridized at 50°C with ^{32}P-ATP end-labeled DNA oligonucleotide probes complementary to the corresponding miRNAs (Supplementary Table 2) in buffer containing 5× SSC, 20mM Na_2HPO_4 pH 7, 7% SDS, 2× Denhardt's Solution and 1 mg of sheared, denatured salmon sperm DNA (Sigma). The blots were stripped and re-probed with a ^{32}P-ATP end-labeled probe complementary to *U6* (Supplementary Table 2) as a reference gene. Both mouse and zebrafish RNA gel blots for miR-29b in Fig. 1d were performed in technical duplicates. The miR-29a and miR-29c hybridizations in Fig. 1d were performed once. The gel blots and hybridizations in Fig. 5a and Supplementary Fig. 6a were performed in technical triplicates for miR-29b and in technical duplicates for miR-29a and miR-29c.

Reverse transcription and real-time PCR

To quantify miR-29a/b/c levels, 10 ng of total RNA were reverse-transcribed using the TaqMan MicroRNA Reverse Transcription Kit (Applied Biosystems) followed by qPCR using TaqMan Universal Master MixII, no UNG (Applied Biosystems). U6 snRNA was used as a control. The following specific TaqMan miRNA assays (Applied Biosystems) were used: Assay IDs mmu-miR-29b (000413), mmu-miR-29a (002112), mmu-miR-29c (000587), U6 snRNA (001973).

For the detection of *Nrep* and *Gapdh* transcript levels, cDNAs from NPCs and mouse brain were produced with the SuperScript IV Reverse Transcriptase kit using oligo-dT (Invitrogen™) and amplified in TaqMan Universal PCR Master Mix, no AmpErase UNG (Applied Biosystems) using the following TaqMan Gene Expression probes: *Nrep* (assay ID Mm00474047_m1) and *Gapdh* (assay ID Mm99999915_g1).

For the detection of *Nestin*, *Oct4*, *Gapdh*, pri-miR-29a/b-1, pri-miR-29b-2/c and *β-actin* transcript levels, cDNAs were produced with the SuperScript IV Reverse Transcriptase kit using oligo-dT (Invitrogen™) and amplified with *Power* SYBR Green PCR Master Mix (Applied Biosystems) using the primers listed in Supplementary Table 2. The precise P values for all qPCR analyses are listed in Supplementary Table 1.

Small RNA-Seq, RNA-Seq and computational analyses

Both RNA and small RNA Illumina libraries were prepared from 1 μg of total RNA isolated using TRIzol (Invitrogen) from a pool of 100, 72 hpf zebrafish embryos or a population of mouse NPCs, respectively, and were sequenced on an Illumina HiSeq 2500 sequencer generating 50 bp single-end reads. Small RNA sequencing libraries were constructed using the NEBNext Multiplex small RNA library kit (NEB) following manufacturer's instructions. PCR amplified cDNA was size selected on a 6% nondenaturing polyacrylamide gel. RNA sequencing libraries were constructed using the TruSeq kit (Illumina) following manufacturer's instructions.

For the RNA-Seq data analysis, FASTQ files were mapped to the zebrafish genome version danRer7 with TopHat version 2.0.6[59] using the global alignment option and a seed size of 22 without allowing mismatches in this region. Only the unique best hits were kept. BAM files from technical duplicates were merged and used to generate pileup tracks with BEDTools[60] and UCSC tools[61].

For the small RNA-Seq data analysis, FASTQ files containing the small RNA sequencing data were stripped of adapter sequences using the Reaper[62] tool and sequences between 16-24nt were selected and mapped using a command-line variant of chimiRa[63] to known microRNA precursors from miRBase (v21)[64]. Each sequence was assigned to a mature miRNA according to which arm (5p or 3p) of the precursor was matched and to the known mature sequence. In cases where a single sequence mapped to multiple possible loci, a random locus was assigned each time. These raw counts were assembled into a table across all microRNAs from all samples (R/BioConductor) and normalized using DESeq2[65]. Differential expression was performed using the negative binomial Wald test (DESeq2) for differences between mutant versus wild type. All derived P-values for differential expression were multiple testing corrected using the Benjamini-Hochberg method.

The trimming and tailing analyses were performed by examining the fraction of nucleotides from the annotated microRNA from miRBase that was covered by each mapped read from each sample. The mean of these values was used to compute differential trimming using a paired t-test on the derived values between 18-23nt (trimming) and 23-28nt (tailing) between mutant and wild type samples for miR-29b.

In situ hybridization on mouse and zebrafish brain sections

Brains from wild type and *Nrep*$^{miR\text{-}29scr}$ three-month old mice were harvested, embedded in OCT (Sakura), and frozen on dry ice (wild type animals n=4, *Nrep*$^{miR\text{-}29scr}$ animals n=4). Zebrafish brains

from six- to eight-month old animals of mixed sex were initially fixed in 4% paraformaldehyde overnight followed by two hour Dietrich's fixative prior to standard paraffin embedding and sectioning.

miRNA *in situ* hybridization on 8 micron sections was performed using 3'-DIG labelled LNA probes for miR-29a, miR-29b, and miR-29c (Exiqon). Scrambled LNA-probes were used as negative hybridization controls. Briefly, sections were fixed in 4% paraformaldehyde, digested with proteinase K for 6 minutes, acetylated, and hybridized with the corresponding probes in 50% formamide, 5× SSC, 5× Denhardt's solution, 500 µg/ml salmon sperm DNA, and 250 µg/ml tRNA overnight at 48°C. Post-hybridization washes were performed in 50% formamide and 2× SSC at 48°C followed by 2× SSC at ambient temperature then the sections were blocked in 10% sheep serum and incubated overnight with the anti-DIG-AP antibody (Roche; at 1:1000) at 4°C. Signal detection was done using NBT/BCIP substrate (Roche).

To generate probes for *in situ* hybridization*, libra* and *Nrep* fragments were amplified by PCR using Phusion DNA polymerase (Thermo Scientific) using the primers listed in Supplementary Table 2, and the PCR products were subcloned into the pGEM-T Vector System (Promega) and confirmed by sequencing. DIG-labeled RNA probes were generated by linearizing the pGEM-T vector and *in vitro* transcribing the probes with T7 or SP6 RNA Polymerase using DIG-RNA Labeling Kit (Roche).

Nrep and *libra in situ* hybridizations were done identically to the miR-29 *in situs* except that the hybridization and wash temperature was 56.5°C.

Mouse brain histology

Brains from wild type and $Nrep^{miR\text{-}29scr}$ three-month old mice were harvested (wild type animals n=4, $Nrep^{miR\text{-}29scr}$ animals n=4), embedded in OCT (Sakura), frozen on dry ice, sectioned at 12 micron, air-dried and fixed in 4% paraformaldehyde. Hematoxylin (Sigma) and eosin (Sigma) staining was performed using standard protocols.

Zebrafish behavioral tests

Six- to eight-month old, mixed sex wild type and mutant zebrafish of the AB genetic background were used for all behavioral studies. Zebrafish were transported from the aquarium into the testing room in their housing tanks one hour before the behavioral analysis in order to acclimate them. The

illumination and temperature of the testing room were similar to the aquarium. The experiments were performed between 10:00 and 15:00 to minimize circadian differences. The age, sex, and size of fish from the different genotypes were matched for each experiment. All recordings were performed using a web camera (HP HD 2300 Webcam) coupled to QuickTime software. Videos were analyzed using the Noldus EthoVision XT8 video-tracking system to obtain the different behavioral endpoints.

Novel tank diving test

Explorative and anxiety-like behavior was measured in the novel tank test as previously described[32]. Single fish were placed into the novel tank and filmed from the side in a 5-minute experiment. We measured the amount of time spent in the bottom, middle and top thirds of the tank, the number of top entries, the total distance swum, the time spent freezing, and the absolute angular velocity using Noldus EthoVision XT8. The novel tank test was performed in three independent experiments.

Aggression test

Aggression was measured using the mirror test as previously described[31]. Individual fish were placed into this setup and recorded from the top for five minutes. The time spent in aggressive display was manually quantified using LabWatcher software (ViewPoint) by an observer blind to the genotype of the fish. Aggressive interactions were scored as time spent biting or pushing against the mirror or thrashing the caudal fin[66]. The aggression test was performed in one experiment.

Novel-object boldness test

Novel-object boldness was measured as previously described[31]. Single fish were placed into a large tank and were recorded from above for five minutes. The novel object was represented by a 15 ml Falcon tube filled with yellow modelling clay. The time spent within one body length of the object and the total distance swum were measured using Noldus EthoVision XT8. The novel-object boldness test was performed in one experiment.

Statistical Analyses

Behavioral data were exported as Excel files (Microsoft) and statistical analyses were carried out in GraphPad Prism6. One-way ANOVA followed by Dunnett's *post hoc* was performed except for latency to approach the novel object where Kruskal-Wallis and Dunn's multiple comparisons *post hoc* tests were performed $*P < 0.05$, $**P < 0.01$, $***P < 0.001$. Data are presented as scatter plots showing the mean with standard error of the mean. The behaviors were assessed for normality

using D'Agostino & Pearson/Shapiro-Wilk normality tests. The variance between the groups was tested using Brown-Forsythe/Barlett's tests. The number of animals tested is denoted by *n* and ranged from 12 to 15 animals per experiment. The precise P values for all behavioral analyses are listed in S2.

Mouse behavioral tests

Adult male 12-16-week-old wild type and homozygous *Nrep*$^{miR\text{-}29scr}$ mice in a C57BL/6N background used for the behavioral tests were housed in groups (2-5 per cage) in a temperature-controlled room (21 $\pm$ 1 °C), relative humidity 50-60%, under a 12-h light/dark cycle (lights on at 7:00 AM), with free access to food and water.

Rotarod test

The rotarod test is used to assess motor coordination and balance in rodents[37]. Mice had to maintain their balance on a 3-cm diameter rod rotating at an accelerating speed of 4 to 40 rpm during 300 s (model 47600; Ugo Basile). To familiarize the mice with the apparatus, mice underwent a training session consisting of three, 60 s trials in which the rod was kept stationary for the first trial and held at 4 rpm for the last two trials. On the following day and for four consecutive days, mice were tested three times per day with an inter-trial interval of 20 min to evaluate motor learning. A maximum of three mice were placed on the rod at the same time. The latency to fall from the rotating rod was recorded in each trial. If a mouse was passively rotating on the rod (i.e. clinging) the number of passive rotations was counted (Supplementary Table 3). For each day, data were expressed as mean latency to fall subtracting one second for each passive rotation[38]. Mice from both genotypes were pseudo-randomly assigned to rotarod sessions. Order of testing was maintained across trials and daily sessions to ensure constant time of testing and inter-trial intervals. Investigators were blinded to the genotype of the mice, both during the rotarod test and the off-line scoring.

Fear conditioning

Fear conditioning was assessed using an automated system (Freeze Monitor, San Diego Instruments, CA, USA) as previously described[40]. Conditioning session (day 1): the mouse was placed in the test chamber (65 lux, 25 x 25 x 19 cm) and allowed to explore it for 9 min. The animal was exposed to a light+tone (140 lux, 92 dB) conditioned stimulus (CS) presented for 20 seconds and paired during the last seconds with two mild (1 second, 0.4mA) foot-shocks (unconditioned

stimulus, US), one after 4 and one after 6.5 min from the beginning of the session. Context session (day 2): after 24 h the mouse was returned to the same testing chamber and its activity was recorded for 6 min during which no cues or shocks were presented. Cue session (day 2): 4 hours after the context session the chamber was modified to change the context of the testing environment (room lights were turned off, the illumination level inside the chamber was 0 Lux, the position of the chamber was rotated of 90◦, the floor and walls were covered with black PVC plates and a cinnamon scent was sprayed inside the chamber). The mouse was placed in this modified chamber and allowed to explore for 8 min. Two and six min after the beginning of the session, the light+tone cues were presented for 2 min (Supplementary Fig. 5a)[40]. In all sessions, freezing behavior was operationally defined as the absence of any movement, except for respiratory ones. An animal is considered to be "freezing" when there are fewer than three beam breaks in 5 sec, recorded automatically in bouts/episodes of freezing of 2 sec minimum[67] and expressed as the percentage of freezing ((100 x the number of freezing episodes)/total number of possible freezing episodes).

Morris Water Maze

Mice were trained to locate a submerged platform (diameter 10 cm) in a water tank (diameter 1 m, temperature 26-28°C) by swimming and relying on external visible cues. Five-day procedure: *Familiarization (day 1),* the mouse was placed on a visible platform and then allowed to swim freely. Then, in three consecutive trials, mice were inserted in the maze from three different starting points. If the mouse did not reach the platform in 60 sec it was guided to the platform. Latency to reach the visible platform was measured; *Training (days 2-4),* the mouse was placed in different maze quadrants randomly. The latency to reach a *hidden platform* (positioned in the Correct quadrant) was measured in three trials per session for two sessions/day (1 hour between sessions) with a cut off of 60 sec. Swim speed, time, distance and path were measured using a videotracking system (Viewpoint, France); *Test (day 5),* the last session of training was followed by a *probe trial.* The hidden platform was removed, the mouse was placed in the center of the pool and the time spent in each quadrant (expressed as % of the total time) and the number of annulus crossings was measured for 60 sec[68].

Elevated plus-maze

This test is used to assess anxiety-related behaviors in mice. The apparatus was made of dark grey PVC and consisted of two open arms (30x5 cm with 0.3cm ledges), two enclosed arms (30x5 cm with 13 cm high side- and end-walls) and a connecting central platform (cm 5x5). The maze was raised to 50 cm above the floor. The illumination level in the maze was approximately 30 lux in

the enclosed arms and 80 lux in the open arms. After 30 min of acclimation to the testing room, animals were placed in the central platform facing one of the open arms. During the 5-min test session the following parameters were scored manually from digital video-recordings: number of open- and closed-arm entries (all four paws crossing) and time spent in open and closed arms[40].

Statistical analysis and Reproducibility

The number of mice necessary for the behavioral experiments was chosen so that, with an anticipated effect size (Cohen's d) ≥ 0.8, alpha=0.05 (2-sided) and power (1-beta) = 0.80, the sample size was at least 7 mice in each experimental group (Power analysis, G*Power 3.1).

All data were analyzed by repeated measures ANOVA with Genotype as between-subjects factor and Day Time or Quadrant as within-subjects factor (with the exception of the Elevated plus maze frequency of total arm entries and % time spent in open arms that were analyzed by t-tests) using the StatView 5.0 PowerPC (SAS Institute Inc., Cary, NC, USA) and Prism 5.0a (GraphPad Software Inc., La Jolla, CA, USA) software packages. *Post hoc* analyses (Bonferroni corrected comparisons) were performed when reputed informative. In all tests, the variance between groups tested with the Bartlett's test for Homogeneity of Variance was not significantly different. The level of significance was set at $P < 0.05$. Data are presented as individual data points and mean ±SEM. Detailed statistical results for all behavioral analyses are listed in Supplementary Table 1. Each mouse participated in each behavioural test one time. If needed, a behavioural test was performed several times with small batches of mice to insure proper mouse handling and appropriate age.

Cell line procedures

ESC growth

Mouse ESCs (male E14 mouse stem cell line[69,70] provided by E. Heard, Institut Curie) were cultured on gelatin-coated flasks in mESC media containing DMEM (Gibco) supplemented with 15% FBS (Gibco), 100 mM nonessential amino acids (Sigma), 0.1 mM 2-mercaptoethanol (Gibco), 1 mM L-Glutamine (Gibco), 1 mM sodium pyruvate (Gibco) and 10 ng/mL leukemia inhibitory factor (LIF, Miltenyi). After two passages, cells were collected and frozen for RNA extraction. All cells tested negative for mycoplasma contamination.

ESC differentiation

mESC E14 cells were differentiated to neural progenitor cells (NPCs) as previously described [41,42] with minor modifications. Briefly, cells were cultured in serum-free N2B27 media (Neurobasal media (Gibco):DMEM/F-12 GlutaMAX (Gibco) 1:1, 1× N2 supplement (Gibco), 0.5× B27 supplement (Gibco), 2 mM L-glutamine (Gibco), 100 μM 2-mercaptoethanol (Gibco), for 7 d, followed by the formation of neural spheres in N2B27 supplemented with EGF and FGF (10 ng/mL, Peprotech). Four-day-old spheres were expanded to NPCs on gelatin-coated culture dishes in N2B27 media supplemented with EGF and FGF (10 ng/mL) for four passages to purify the NPC cultures after which they were collected and immediately frozen for RNA extraction.

Generation of Nrep$^{miR\text{-}29scr}$ ESC knock-in lines

To create the donor plasmid, a 912 bp *Nrep* fragment flanking the miR-29b binding site was PCR amplified with primers EcoRV-*Nrep* and Nru1-*Nrep* (Supplementary Table 2) and cloned into the EcoRV-Nru1 linearized pBR322 plasmid using the NEBuilder HiFi DNA assembly cloning kit (NEB). The *Nrep* miR-29b site and the PAM sequence corresponding to the sgRNA Guide5 target site (Supplementary Fig. 4) were mutagenized simultaneously using the QuickChange Lightning site-directed mutagenesis kit (Agilent) as recommended by the manufacturer with the primers *Nrep*-sdm-For and *Nrep*-sdm-Rev (Supplementary Table 2). The mutagenized *Nrep*-pBR322 plasmid was verified by DNA sequencing.

mESCs were transfected with Lipofectamine 2000 (Invitrogen) following the manufacturer's guidelines with 400ng of the *Nrep*mutpBR322 plasmid and 400ng of the pSpCas9(BB)-2A-Puro (pX459) CRISPR/Cas9 vector[71] containing the sgRNA Guide5 sequence (Supplementary Table 2). Cells were selected for 48 hours in 1.0 μg/mL of puromycin (Sigma) and cultured an additional 10 days on gelatin-coated 10 cm plates (Falcon) in mESC media (defined above) before individual clones were selected and submitted to PCR selection and DNA sequencing to verify the integrity of both the *Nrep* wild type and *Nrep$^{miR\text{-}29scr}$* loci. The presence of the scrambled sequence was scored by PCR using the genotyping primers listed in Supplementary Table 2 followed by BamH1 digestion of the amplified DNA region.

Data Availability and Accession Code Statement

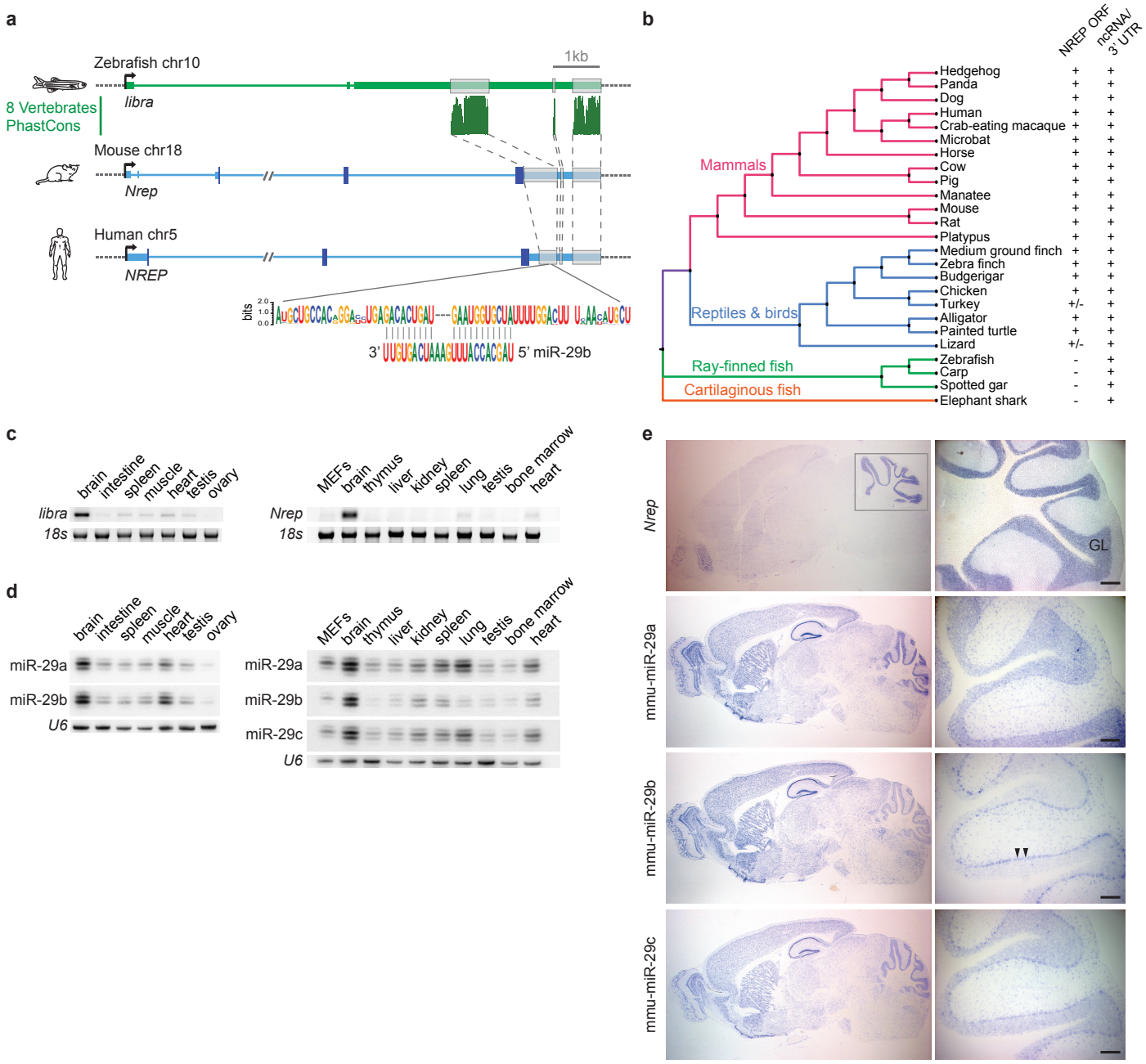
The high-throughput sequencing data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus and are accessible through GEO Series accession number GSE98707 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE98707).

Source data for Figure 2, 4, and 5 and Supplementary Figures 3, 4, 5 and 6 are available with the paper online as Supplementary Data Set 2. Computer code used to analyze small RNA sequencing is available upon request from A.J.E. The *Nrep*$^{miR\text{-}29scr}$, *libra*^{del} and *libra*^{inv} alleles will be made available upon request.

METHODS-ONLY REFERENCES

48. Hwang, H.W., Wentzel, E.A. & Mendell, J.T. A hexanucleotide element directs microRNA nuclear import. *Science* **315**, 97-100 (2007).
49. Griffiths-Jones, S., Saini, H.K., van Dongen, S. & Enright, A.J. miRBase: tools for microRNA genomics. *Nucleic Acids Res* **36**, D154-8 (2008).
50. Auer, T.O., Duroure, K., De Cian, A., Concordet, J.P. & Del Bene, F. Highly efficient CRISPR/Cas9-mediated knock-in in zebrafish by homology-independent DNA repair. *Genome Res* **24**, 142-53 (2014).
51. Hwang, W.Y. et al. Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat Biotechnol* **31**, 227-9 (2013).
52. Kimmel, C.B., Ballard, W.W., Kimmel, S.R., Ullmann, B. & Schilling, T.F. Stages of embryonic development of the zebrafish. *Dev Dyn* **203**, 253-310 (1995).
53. Yang, H. et al. One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. *Cell* **154**, 1370-9 (2013).
54. Goujon, M. et al. A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic Acids Res* **38**, W695-9 (2010).
55. Sievers, F. et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**, 539 (2011).
56. McWilliam, H. et al. Analysis Tool Web Services from the EMBL-EBI. *Nucleic Acids Res* **41**, W597-600 (2013).
57. Crooks, G.E., Hon, G., Chandonia, J.M. & Brenner, S.E. WebLogo: a sequence logo generator. *Genome Res* **14**, 1188-90 (2004).
58. Pall, G.S. & Hamilton, A.J. Improved northern blot method for enhanced detection of small RNA. *Nat Protoc* **3**, 1077-84 (2008).
59. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S.L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**, R25 (2009).
60. Quinlan, A.R. & Hall, I.M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841-2 (2010).
61. Kent, W.J., Zweig, A.S., Barber, G., Hinrichs, A.S. & Karolchik, D. BigWig and BigBed: enabling browsing of large distributed datasets. *Bioinformatics* **26**, 2204-7 (2010).
62. Davis, M.P., van Dongen, S., Abreu-Goodger, C., Bartonicek, N. & Enright, A.J. Kraken: a set of tools for quality control and analysis of high-throughput sequence data. *Methods* **63**, 41-9 (2013).
63. Vitsios, D.M. & Enright, A.J. Chimira: analysis of small RNA sequencing data and microRNA modifications. *Bioinformatics* **31**, 3365-7 (2015).
64. Kozomara, A. & Griffiths-Jones, S. miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res* **42**, D68-73 (2014).

65. Love, M.I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).
66. Gerlai, R., Lahav, M., Guo, S. & Rosenthal, A. Drinks like a fish: zebra fish (Danio rerio) as a behavior genetic model to study alcohol effects. *Pharmacol Biochem Behav* **67**, 773-82 (2000).
67. Curzon, P., Rustay, N.R. & Browman, K.E. Cued and Contextual Fear Conditioning for Rodents. in *Methods of Behavior Analysis in Neuroscience* (eds. nd & Buccafusco, J.J.) (Boca Raton (FL), 2009).
68. Simon, M.M. et al. A comparative phenotypic and genomic analysis of C57BL/6J and C57BL/6N mouse strains. *Genome Biol* **14**, R82 (2013).
69. Nora, E.P. et al. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* **485**, 381-5 (2012).
70. Schulz, E.G. et al. The two active X chromosomes in female ESCs block exit from the pluripotent state by modulating the ESC signaling network. *Cell Stem Cell* **14**, 203-16 (2014).
71. Ran, F.A. et al. Genome engineering using the CRISPR-Cas9 system. *Nat Protoc* **8**, 2281-308 (2013).


a
Zebrafish chr10
1kb
8 Vertebrates PhastCons
libra
Mouse chr18
Nrep
Human chr5
NREP
bits
3' UUCUCACUAAAGUUUACCACGAU 5' miR-29b
b
NREP ORF
ncRNA/ 3' UTR
Hedgehog + +
Panda + +
Dog + +
Human + +
Crab-eating macaque + +
Microbat + +
Horse + +
Cow + +
Pig + +
Manatee + +
Mouse + +
Rat + +
Platypus + +
Medium ground finch + +
Zebra finch + +
Budgerigar + +
Chicken + +
Turkey +/- +
Alligator + +
Painted turtle + +
Lizard +/- +
Zebrafish - +
Carp - +
Spotted gar - +
Elephant shark - +
Mammals
Reptiles & birds
Ray-finned fish
Cartilaginous fish
c
brain intestine spleen muscle heart testis ovary
libra
18s
MEFs brain thymus liver kidney spleen lung testis bone marrow heart
Nrep
18s
d
brain intestine spleen muscle heart testis ovary
miR-29a
miR-29b
U6
MEFs brain thymus liver kidney spleen lung testis bone marrow heart
miR-29a
miR-29b
miR-29c
U6
e
Nrep
GL
mmu-miR-29a
mmu-miR-29b
mmu-miR-29c

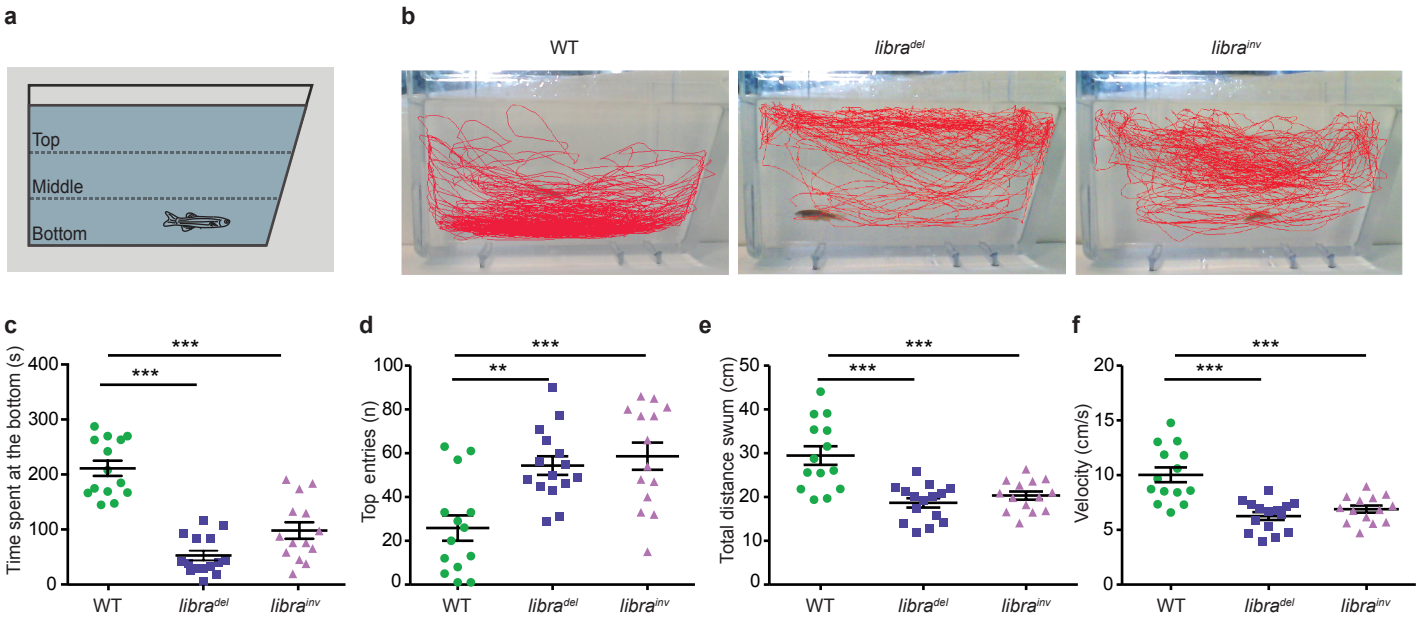
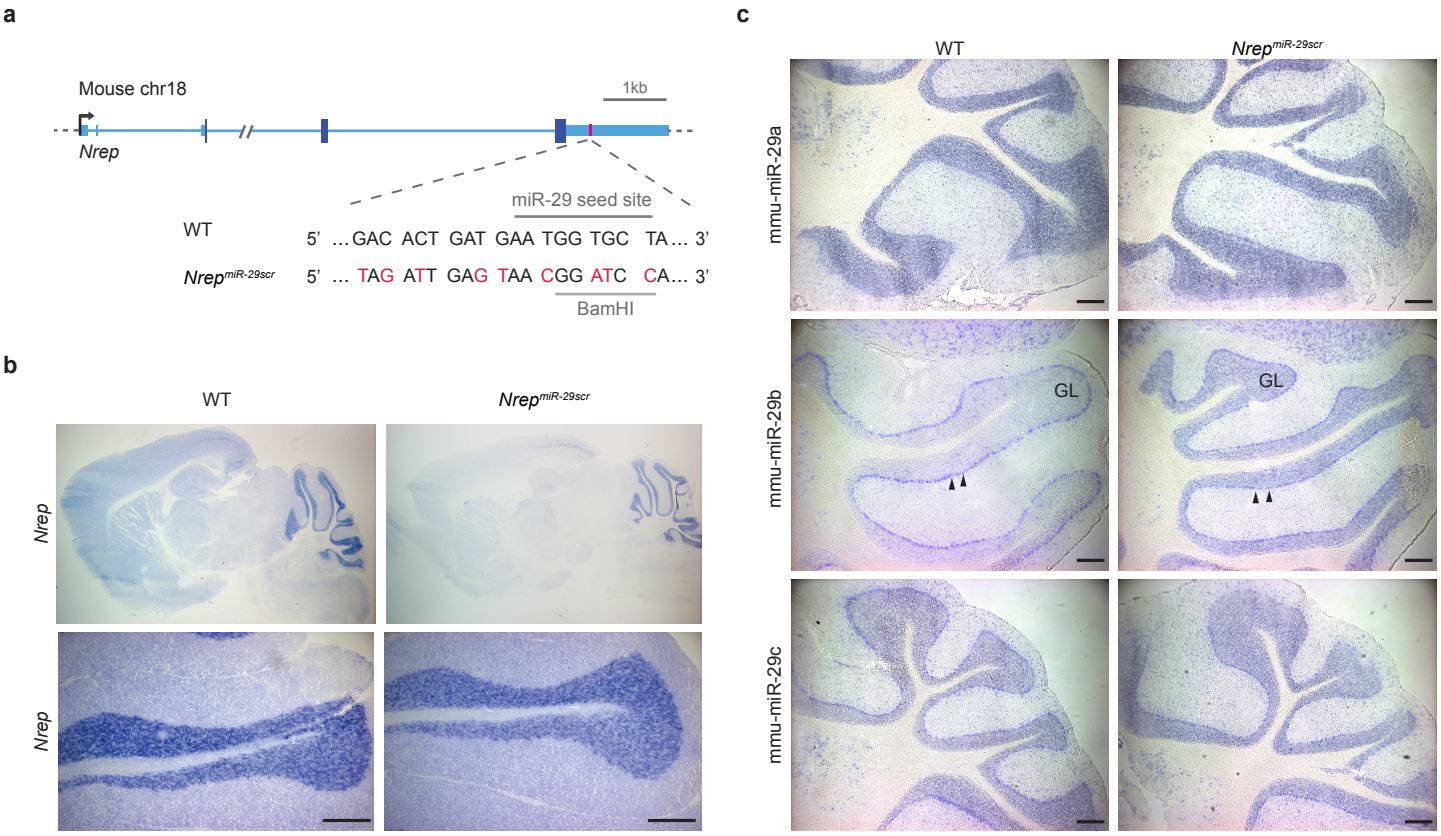
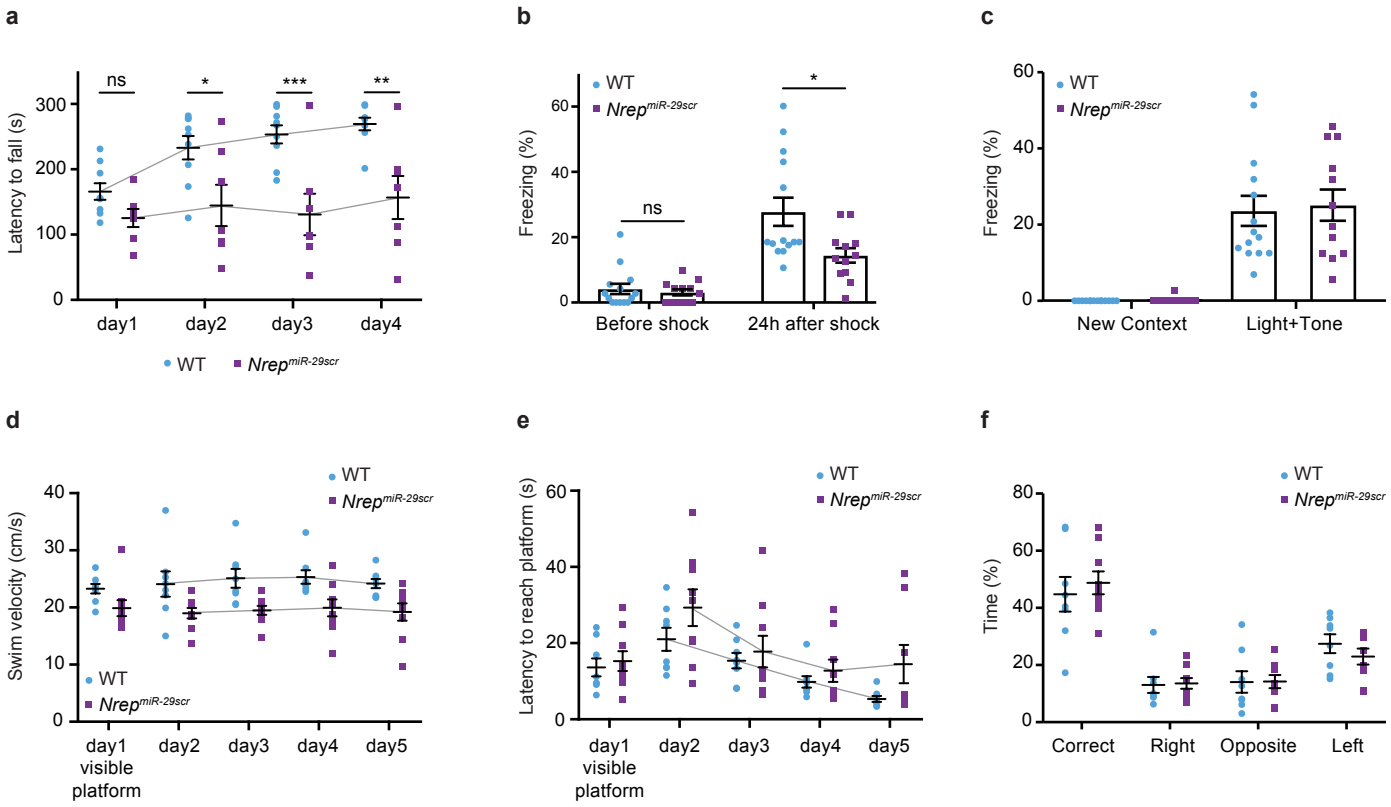
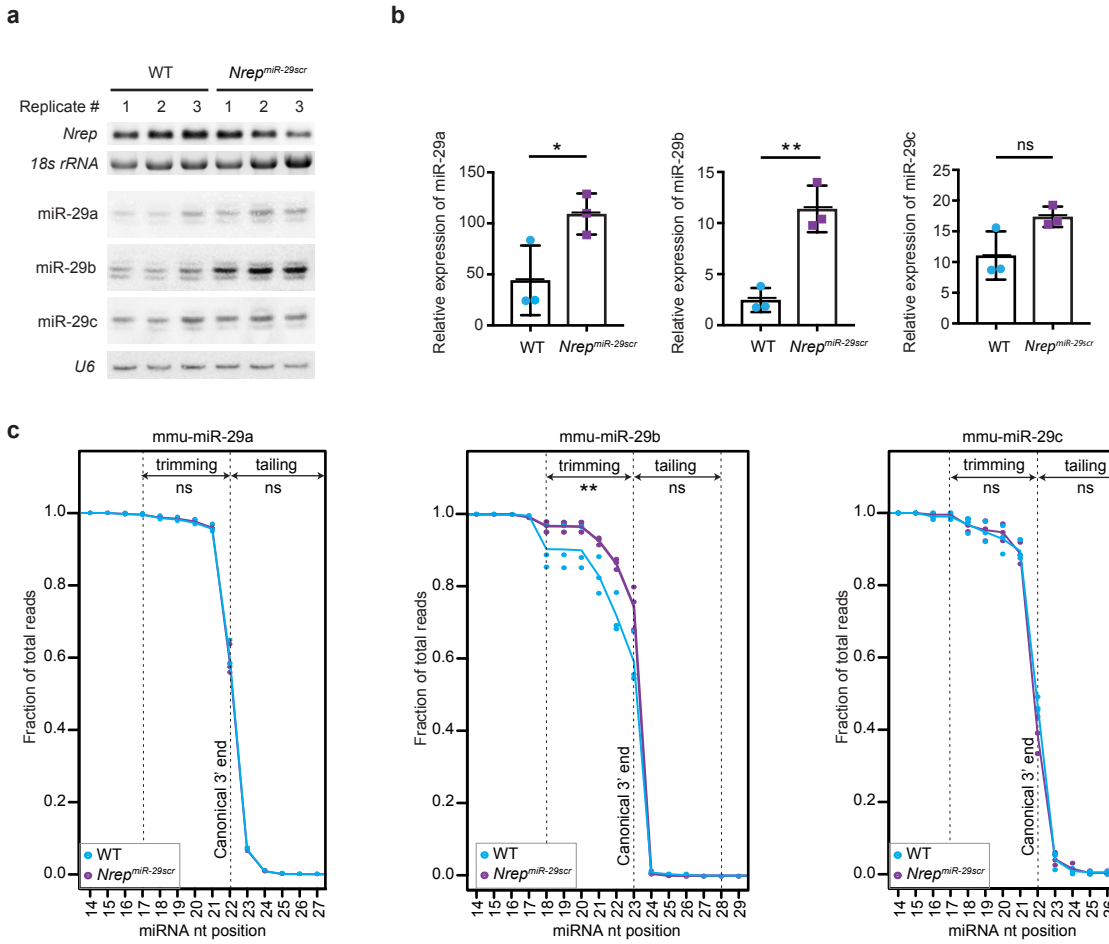

a
Top
Middle
Bottom
b
WT
libradel
librainv
c
Time spent at the bottom (s)
0
100
200
300
400
***
***
WT
libradel
librainv
d
Top entries (n)
0
20
40
60
80
100
***
**
WT
libradel
librainv
e
Total distance swum (cm)
0
10
20
30
40
50
***
***
WT
libradel
librainv
f
Velocity (cm/s)
0
5
10
15
20
***
***
WT
libradel
librainv

Figure 3

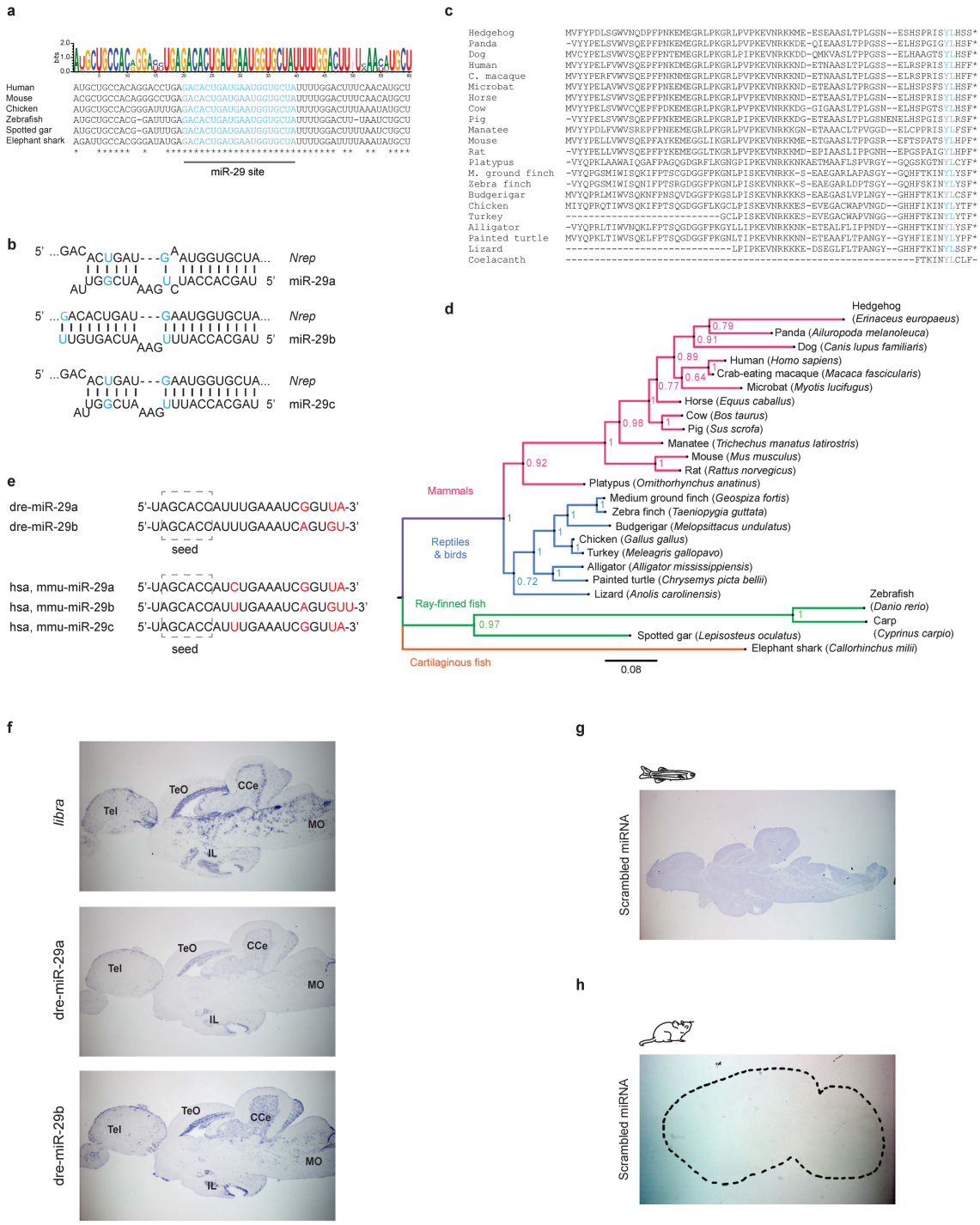

a
Mouse chr18
1kb
Nrep
miR-29 seed site
WT
5' ...GAC ACT GAT GAA TGG TGC TA... 3'
Nrep^miR-29scr
5' ... TAG ATT GAG TAA CGG ATC CA... 3'
BamHI
b
WT
Nrep^miR-29scr
Nrep
Nrep
c
WT
Nrep^miR-29scr
mmu-miR-29a
mmu-miR-29b
mmu-miR-29c
GL
GL


a
Latency to fall (s)
ns
*
***
**
day1
day2
day3
day4
WT
Nrep miR-29scr
b
Freezing (%)
WT
Nrep miR-29scr
ns
*
Before shock
24h after shock
c
Freezing (%)
WT
Nrep miR-29scr
New Context
Light+Tone
d
Swim velocity (cm/s)
WT
Nrep miR-29scr
day1 visible platform
day2
day3
day4
day5
e
Latency to reach platform (s)
WT
Nrep miR-29scr
day1 visible platform
day2
day3
day4
day5
f
Time (%)
WT
Nrep miR-29scr
Correct
Right
Opposite
Left


a
WT
Nrep miR-29scr
Replicate # 1 2 3 1 2 3
Nrep
18s rRNA
miR-29a
miR-29b
miR-29c
U6
b
Relative expression of miR-29a
Relative expression of miR-29b
Relative expression of miR-29c
*
**
ns
WT
Nrep miR-29scr
c
mmu-miR-29a
mmu-miR-29b
mmu-miR-29c
trimming
tailing
ns
**
Fraction of total reads
Canonical 3' end
miRNA nt position

INTEGRATED SUPPLEMENTARY FIGURE TEMPLATE

Supplementary Fig. 1

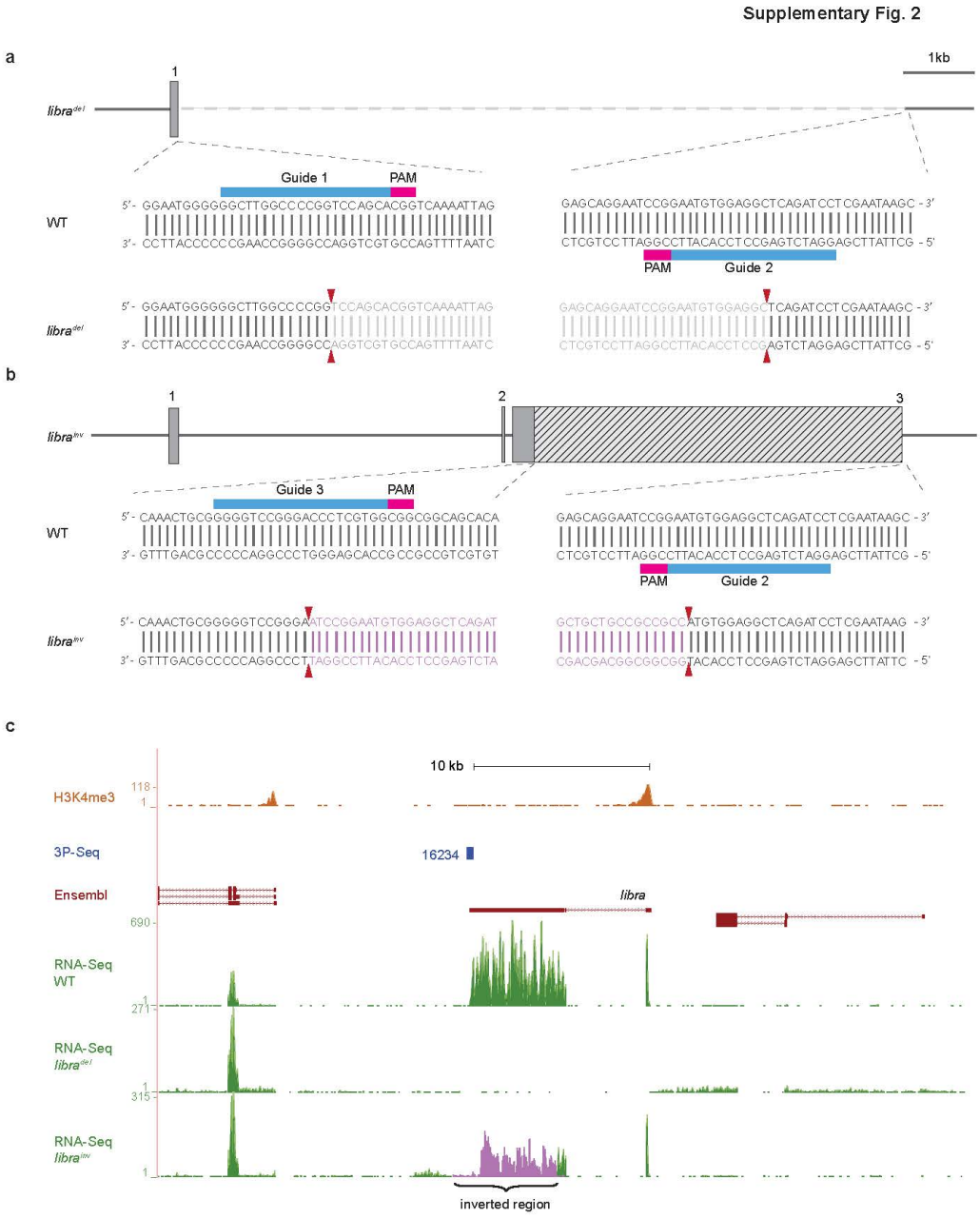


Supplementary Figure 1

Conservation of the *libra*, *Nrep* and miR-29 sequences.

(a) Conservation of the miR-29 site and its flanking sequences. The sequence logo based on 30 homologous sequences is shown above representative examples from the indicated species. Asterisks indicate bases conserved in all of the representative examples. The miR-29 site is in blue and underlined. (b) Predicted *Nrep* pairing with the individual mouse miR-29 family members. Watson-Crick paired nucleotides are in black and connected by vertical lines, whereas wobbled paired nucleotides are in blue. (c) Multiple alignment of the NREP ORF across vertebrates. Amino acids conserved in all species are in blue. An asterisk indicates a putative stop codon. (d) Unrooted consensus Bayesian phylogenetic tree of zebrafish *libra* and its homologs in 24 vertebrate species. Branch lengths represent the number of substitutions per site as indicated in the scale bar. Posterior probabilities for all branch splits are displayed in the nodes. (e) The miR-29 family members of zebrafish (a and b) and human and mouse (a, b and c). The seed sequence is boxed, bases differing among the individual miR-29s are in red. (f) *libra*, miR-29a and miR-29b show partially overlapping expression detected by *in situ* hybridization on adult zebrafish brain sections. Representative sections for each probe are shown (n=6 wild type animals). Telencephalon (Tel), Tectum opticum (TeO), Corpus Cerebelli (CCe), Medulla Oblongata (MO), Inferior Lobe (IL). (g) Control *in situ* hybridization on zebrafish adult brain sections using scrambled miRNA probes. (h) Control *in situ* hybridization on mouse adult brain sections using scrambled miRNA probes. The brain section is outlined with a dashed line. Representative sections for each probe are shown (wild type zebrafish n=6 and mouse n=4).

Click inside this box and **insert a single image for Supplementary Figure 2**.



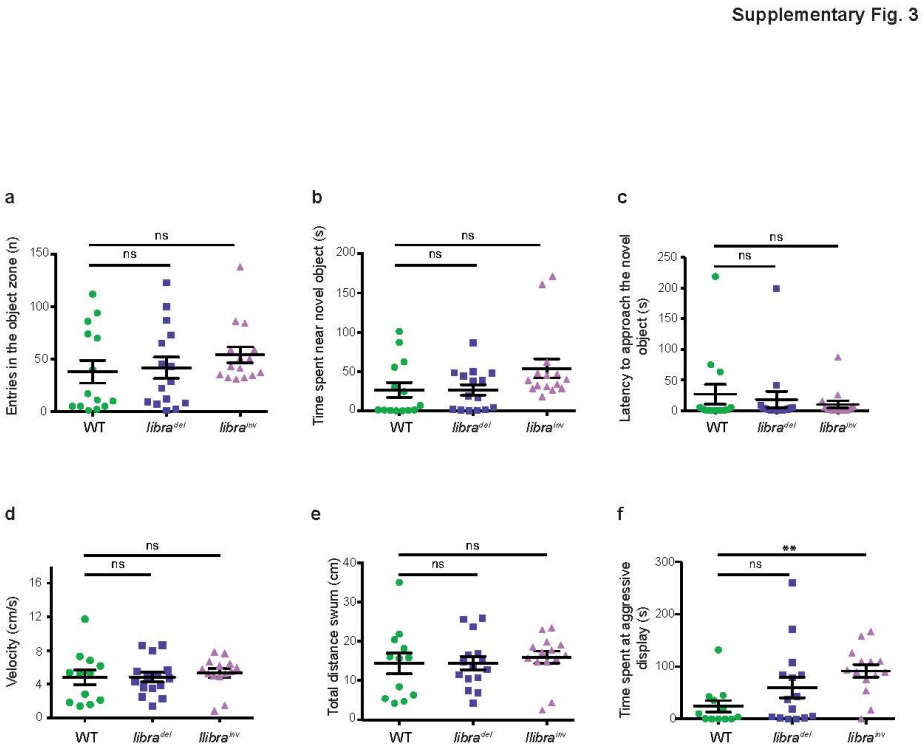
Supplementary Figure 2

Generation and validation of the *libra*^{del} and *libra*^{inv} zebrafish mutants.

(a) The *libra*^{del} zebrafish locus showing the positions of the sgRNA Guide 1 and Guide 2 used to generate the *libra*^{del} mutant. The nucleotides defining the borders of the deleted genomic sequence block are shown in gray and delineated by red arrowheads. The short guide and PAM sequences are indicated with blue and pink blocks, respectively. (b) The *libra*^{inv} zebrafish locus showing the positions of the sgRNA Guide 2 and Guide 3 used to generate the *libra*^{inv} mutant and the inverted portion of exon 3 (hashed

block). The nucleotides defining the borders of the inverted part of the *libra* transcript in the *libra*^{inv} mutant are indicated in lavender and delineated by red arrowheads. The short guide and PAM sequences are indicated with blue and pink blocks, respectively. (c) Generation of the *libra*^{del} and *libra*^{inv} alleles was confirmed by RNA-Seq. The positions of the *libra* locus together with that of the adjacent annotated upstream and downstream loci are indicated, and the corresponding H3K4me3 ChIP-Seq, 3P-Seq and RNA-Seq tracks from wild type, *libra*^{del} and *libra*^{inv} zebrafish are shown. The *y*-axis of the RNA-Seq tracks represents raw read counts. The inverted portion of the *libra* transcript in *libra*^{inv} mutant is indicated in lavender.

Click inside this box and **insert a single image for Supplementary Figure 3**.

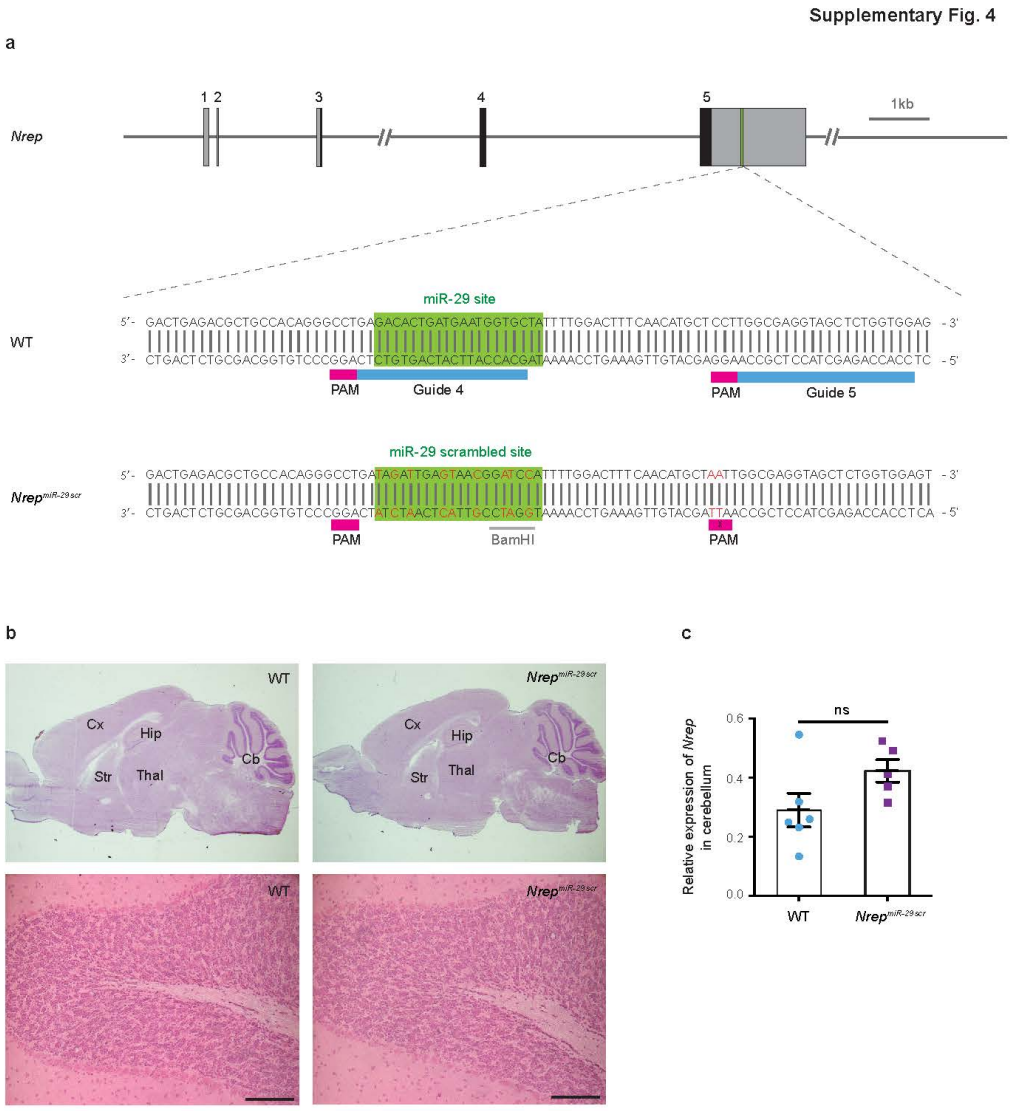


Supplementary Figure 3

Behavior of the *libra*^{del} and *libra*^{inv} zebrafish mutants.

(a) The number (n) of entries, (b) time in seconds (s) spent near the novel object testing boldness and (c) latency to approach the novel object in seconds (s) of wild type and *libra* mutants. Each dot represents an individual adult fish from one experiment; wild type animals n=14, *libra*^{del} animals n=15, *libra*^{inv} animals n=15. (d) Velocity in centimeters per second (cm/s), (e) total distance swum in centimeters (cm) and (f) time in seconds (s) spent at the aggressive display recorded in the aggression test for the indicated genotypes. Each dot represents an individual adult fish from one experiment; wild type animals n=12, *libra*^{del} animals n=15, *libra*^{inv} animals n=14, $^{**}P < 0.01$; (ns) not significant. Data are presented as mean ± SEM, One-way ANOVA and Dunnett's multiple comparisons tests were performed for all analyses except for latency to approach the novel object where Kruskal-Wallis and Dunn's multiple comparisons tests were performed. The novel-object boldness and aggression tests were each performed in one experiment. Detailed statistical analyses and source data are provided in Supplementary Table 1 and Supplementary Data Set 2.

Click inside this box and **insert a single image for Supplementary Figure 4**.



Supplementary Figure 4

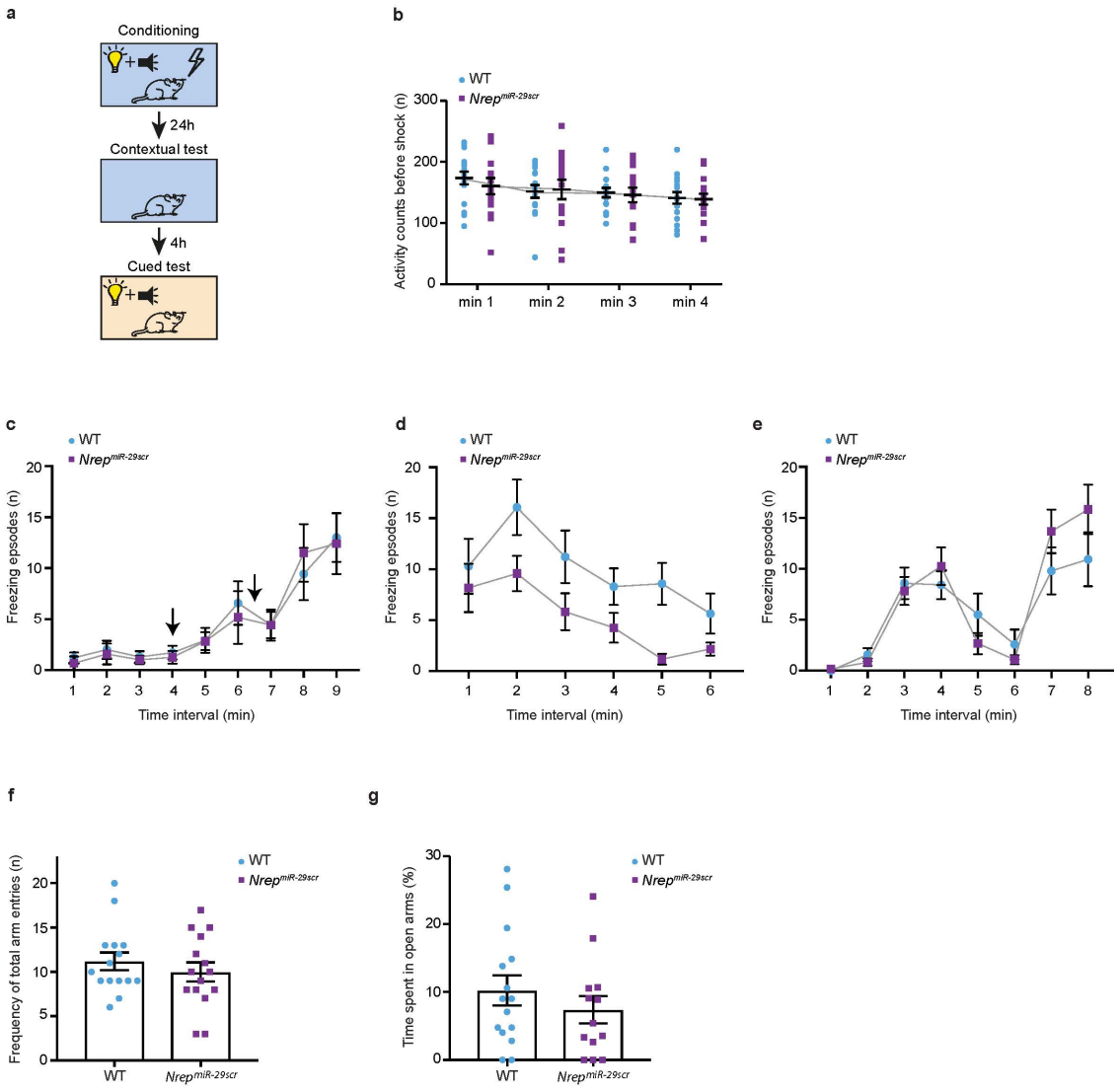
Generation and expression analyses of the mouse *Nrep*$^{miR\text{-}29scr}$ allele.

(a) The mouse *Nrep* locus showing the positions of the sgRNA Guide 4 and Guide 5 used to generate the *Nrep*$^{miR\text{-}29scr}$ mutant mice. The short guide and PAM sequences are indicated with blue and pink blocks, respectively. The wild type and scrambled miR-29 site is highlighted in green, and the scrambled nucleotides are in red. A *BamHI* restriction site was introduced to facilitate mutant screening. (b) Hematoxylin and eosin staining of brain sections from wild type (left) and *Nrep*$^{miR\text{-}29scr}$ (right) mice. Upper panel, whole brain; bottom panel, zoom-in of the cerebellum. Scale bars represent 50 µm. Representative sections for the indicated genotype are shown (wild type animals n=4, *Nrep*$^{miR\text{-}29scr}$ animals n=4). Cortex (Cx), Striatum (Str), Hippocampus (Hip), Thalamus (Thal), Cerebellum (Cb). (c) qRT-PCR analysis of *Nrep* expression in the cerebellum of wild type (blue) and *Nrep*$^{miR\text{-}29scr}$ (purple) mutant mice (wild type animals

n=6, *Nrep*$^{miR\text{-}29scr}$ animals n=5). *Gapdh* was used as a reference gene. Data are presented as mean ± SEM; unpaired t-test: (ns) not significant at 95% confidence level. Detailed statistical analyses and source data for panel c are provided in Supplementary Table 1 and Supplementary Data Set 2.

Click inside this box and **insert a single image for Supplementary Figure 5**.

Supplementary Fig. 5



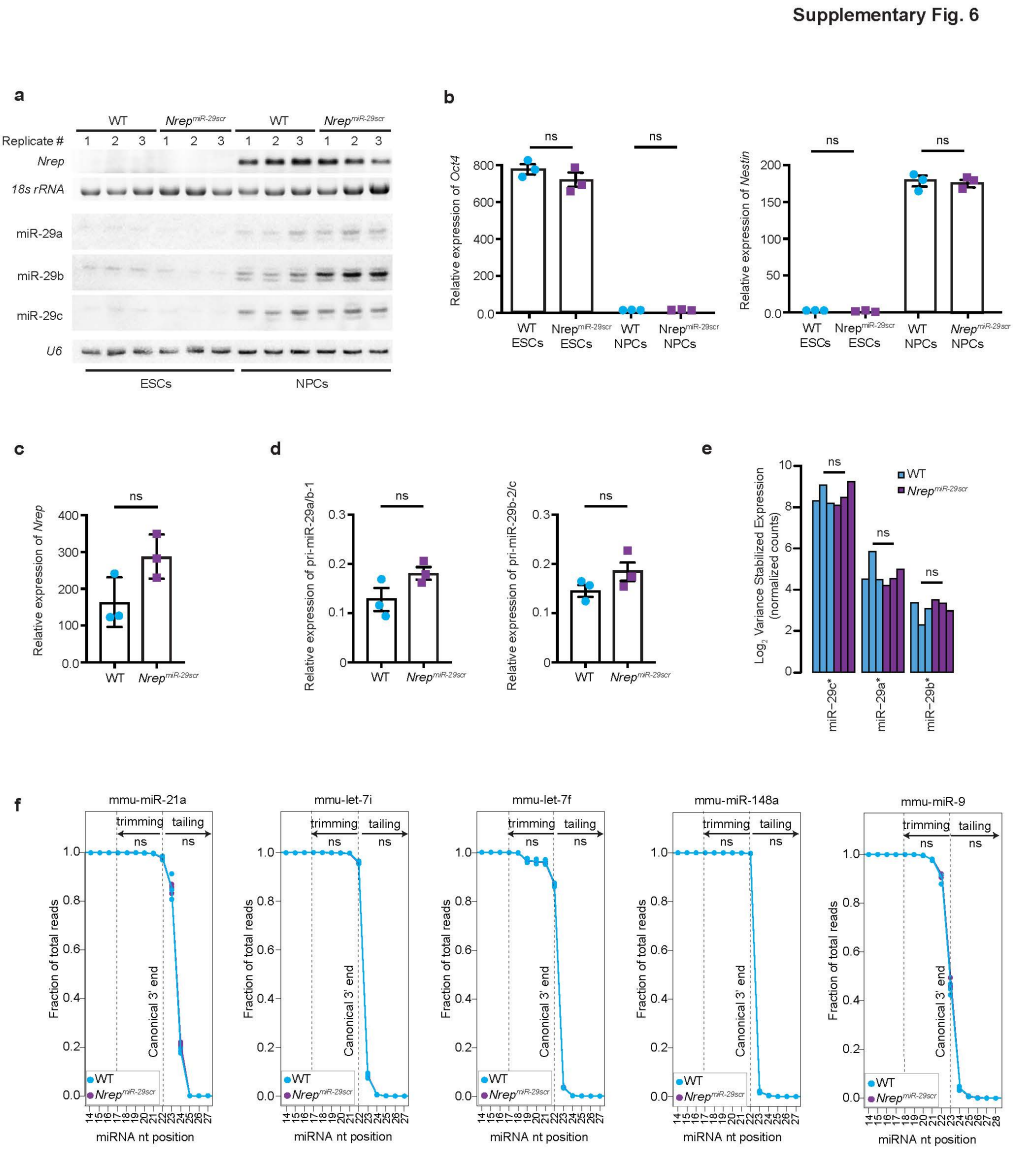
Supplementary Figure 5

Behavior of *Nrep*$^{miR\text{-}29scr}$ mice.

(a) Work flow of the three phases of the contextual and cued fear conditioning test. (b) Number of activity counts before shock in the fear conditioning test; wild type (WT) animals n=16, *Nrep*$^{miR\text{-}29scr}$ animals n=15. Data are presented as individual data points and mean ± SEM. (c-e) Number of freezing episodes of *Nrep*$^{miR\text{-}29scr}$ (purple) and wild type (blue) mice in the fear conditioning test presented in 1 min intervals during (c) the conditioning session (arrows indicate the time of the CS + foot-shock application); (d) the context session; (e) the cue session (CS presented during min 3-4 and 7-8). Wild

type animals n=14, *Nrep*$^{miR\text{-}29scr}$ animals n=12. Data are presented as mean ± SEM. (f-g) Elevated plus maze test to assess anxiety-related behavior in *Nrep*$^{miR\text{-}29scr}$ (purple) and wild type (blue) mice. (f) frequency of total arm entries; wild type animals n=15, *Nrep*$^{miR\text{-}29scr}$ animals n=15, and (g) percent time spent in open arms; wild type animals n=15, *Nrep*$^{miR\text{-}29scr}$ animals n=13. Data are presented as individual data points and mean ± SEM. Detailed statistical analyses and source data for panels b-g are provided in Supplementary Table 1 and Supplementary Data Set 2.

Click inside this box and **insert a single image for Supplementary Figure 6**.



Supplementary Figure 6

Molecular characterization of *Nrep*$^{miR\text{-}29scr}$ cell lines.

(a) *Nrep* and miR-29a, b and c expression in wild type and *Nrep*$^{miR\text{-}29scr}$ ESCs and NPCs detected by RNA blots. The NPC portion of this blot is also shown in Fig. 5a. Three biological replicates are shown for each genotype, technical triplicates were performed for miR-29b and technical duplicates were performed for *Nrep*, miR-29a and miR-29c. *18S* rRNA and *U6* RNAs were used as loading references. Uncropped blot images are shown in Supplementary Data Set 1. (b) qRT-PCR analysis of *Oct4* and *Nestin* expression in wild type and *Nrep*$^{miR\text{-}29scr}$ ESCs and NPCs. Each dot represents an individual biological replicate ESC or NPC population. qRT-PCRs for each biological replicate were performed in technical triplicate. *Gapdh* was

used as a reference gene. Data are presented as mean ± SEM; unpaired t-test: (ns) not significant at 95% confidence level. (c) qRT-PCR analysis of *Nrep* expression in wild type and *Nrep*$^{miR\text{-}29scr}$ NPCs. Each dot represents an individual biological replicate NPC population. qRT-PCRs for each biological replicate were performed in technical triplicate*. Gapdh* was used as a reference gene. Data are presented as mean ± SEM; unpaired t-test: (ns) not significant at 95% confidence level. (d) qRT-PCR analysis of pri-miR-29 expression in wild type and *Nrep*$^{miR\text{-}29scr}$ NPCs. Each dot represents an individual biological replicate NPC population. *β-actin* was used as a reference gene. Data are presented as mean ± SEM; unpaired t-test: (ns) not significant at 95% confidence level (e) Normalized expression of miR-29 star sequences from small RNA sequencing reads. Each bar represents an individual biological replicate NPC population; unpaired t-test: (ns) not significant. (f) The proportion of small RNA sequencing reads with coverage of at least *n* indicated nucleotides (nt) along the length of the top five most abundant NPC mmu-miRNAs (miR-21a, let-7i, let-7f, miR-148a and miR-9). For each miRNA, the canonical length and regions tested for trimming and tailing are delineated. Five nucleotides either side of the canonical length were used to test for significant differences in the mean values over these positions using a paired t-test for difference of means. Three biological replicates are shown for each genotype. (ns) not significant at 95% confidence level. Detailed statistical analyses and source data for panels b-f are provided in Supplementary Table 1 and Supplementary Data Set 2.